

Министерство образования и науки Российской Федерации

Уральский федеральный университет
имени первого Президента России Б. Н. Ельцина

С. М. БОРОДАЧЁВ

**ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ
В ЭКОНОМИКЕ**

Учебное пособие

Екатеринбург

УрФУ

2010

УДК 519.86 (075.8)

ББК 65.054я73

Б83

Бородачёв С. М.

Б83 Имитационное моделирование в экономике : учеб. пособие /

С. М. Бородачёв. – Екатеринбург : УрФУ, 2010. – 81 с.

ISBN 978-5-321-01817-0

Имитационное моделирование – метод анализа и синтеза сложных систем, который используется, когда традиционные математические методы исследования невозможны или неэффективны. Пособие содержит теоретический материал, упражнения по всем разделам курса, лабораторный практикум и задания для курсовых работ.

Предназначено для студентов информационно-математических и экономических специальностей.

Библиогр.: 23 назв.

УДК 519.86 (075.8)

ББК 65.054я73

ISBN 978-5-321-01817-0

©УрФУ, 2010

© Бородачёв С. М., 2010

Оглавление

1. ВВЕДЕНИЕ	5
2. ВЕРОЯТНОСТНО-СТАТИСТИЧЕСКИЕ АСПЕКТЫ	7
ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ	7
2.1. Критерий согласия Колмогорова	8
2.2. Критерий согласия χ^2	9
2.3. Генерация псевдослучайных чисел,	11
равномерно распределённых на отрезке $[0, 1]$	12
2.4. Критерий случайности	12
2.5. Мультипликативный конгруэнтный метод генерации псевдослучайных чисел	14
2.6. Моделирование дискретной случайной величины	15
2.7. Моделирование непрерывной случайной величины	16
с произвольным распределением	16
2.8. Моделирование случайных векторов	17
2.9. Статистический анализ результатов моделирования	20
2.10. Упражнения и лабораторный практикум	21
3. ПАКЕТ <i>MATHCAD</i>	25
Пример: критерий согласия χ^2 и гистограмма	25
3.1. Упражнения и лабораторный практикум	26
4. ПРИМЕРЫ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ	30
4.1. Простая система с влиянием внешней среды	30
4.2. Доверительная оценка погрешности имитационных расчётов	31
4.3. Вычисление определённых и кратных интегралов	31
методом Монте-Карло	31
4.4. Имитационное моделирование инвестиционных рисков	34
Решение в MathCAD	35
4.5. Определение цены опциона методом имитационного моделирования	37
4.6. Моделирование кредитных рисков	38

Экономичный капитал (ЕС – economic capital).....	41
Распределение портфельных потерь.....	42
4.7. Упражнения и лабораторный практикум	43
5. ОБЩИЕ ПРИНЦИПЫ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ МНОГОКОМПОНЕНТНЫХ СИСТЕМ.....	49
5.1. Организация квазипараллелизма просмотром активностей.....	51
5.2. Два способа изменения (протяжки) системного времени	52
5.3. Организация квазипараллелизма транзактным способом	56
5.4. Упражнения и лабораторный практикум	59
6. ИСПЫТАНИЕ И ЭКСПЛУАТАЦИЯ ИМИТАЦИОННЫХ МОДЕЛЕЙ...	60
6.1. Верификация и проверка адекватности модели	60
6.2. Оценка погрешности результирующего показателя имитации	61
из-за различия затравочных чисел генератора псевдослучайных чисел.....	61
6.3. Методы понижения дисперсии.....	62
Антитетический метод	62
Понижение дисперсии при вычислении интегралов.....	63
6.4. Упражнения и лабораторный практикум	64
7. ПРИМЕНЕНИЕ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ К СРАВНЕНИЮ МЕТОДОВ ОЦЕНИВАНИЯ И АНАЛИЗУ ИХ ТОЧНОСТИ....	65
7.1. Упражнения и лабораторный практикум	67
8. ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ	69
БИБЛИОГРАФИЧЕСКИЙ СПИСОК.....	80

1. Введение

Модель – объект-заместитель объекта-оригинала, обеспечивающий изучение некоторых свойств оригинала.

Виды моделирования:

- натурное (модель самолёта из дерева для исследований в аэродинамической трубе; в Англии была построена гидромодель национальной экономики);
- математическое (уравнения, неравенства, которые описывают реальную систему).

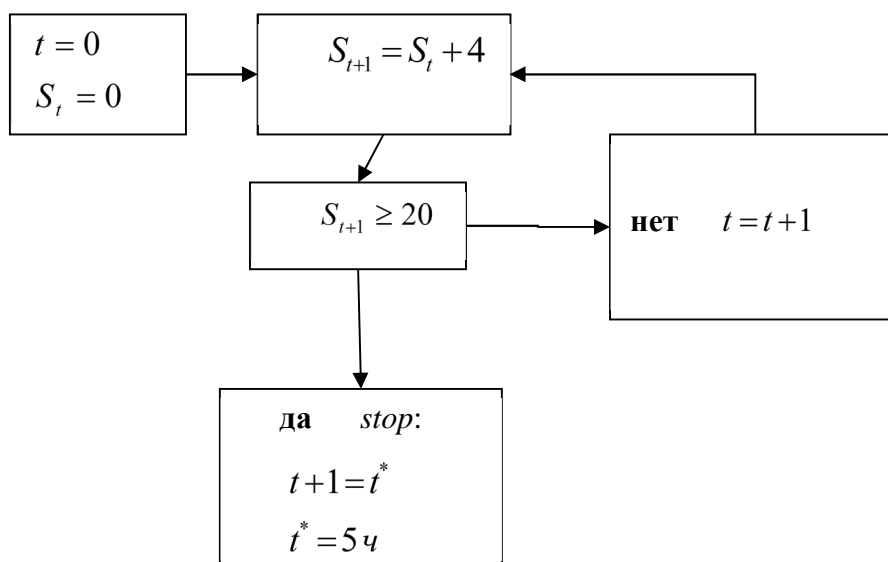
Получение решения на математической модели может быть:

- дедуктивное, т. е. из общих математических свойств, теорем, законов находится решение частной задачи (решение в аналитическом виде);
- имитационное (ускоренное проигрывание процесса на компьютере).

Пример 1.1. Человек прошёл 20 км со скоростью 4 км/ч. Какое время он затратил?

Согласно закону равномерного движения $S = vt$ (математическая модель).

1. Получим дедуктивное решение, разделив обе части на v : $t^* = \frac{S}{v} = \frac{20}{4} = 5$ ч.
2. Имитация: обозначим за S_t путь, пройденный к моменту времени t . За каждый час путь увеличивается на 4 км. Посмотрим, к какому моменту наберется 20 км. Составляем алгоритм (блок-схему), кодируем его на каком-нибудь языке программирования, запускаем программу и находим тот же ответ: $t^* = 5$ ч.



В случае сложной системы построить модель, допускающую дедуктивное решение, бывает трудно или невозможно. Тогда прибегают к имитационному моделированию.

Определение. Имитационное моделирование – серия численных экспериментов, призванных получить оценки степени влияния различных факторов на интересующие нас результаты. Имитационное моделирование с применением случайных чисел называют методом Монте-Карло (город в Монако – княжество между Францией и Италией, – славящийся игорными домами).

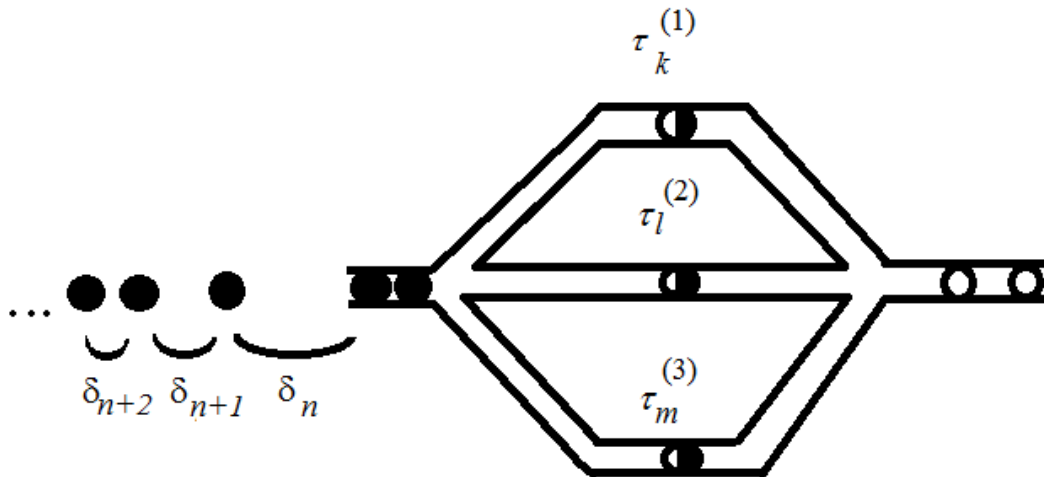
По данным одного интернет-обзора, при решении проблем управления фирмами наиболее часто используются следующие методы:

- имитационное моделирование (29%);
- линейное программирование (21%);
- сетевые методы планирования и управления (14%);
- теория управления запасами (12%);
- нелинейное программирование (8%);
- динамическое программирование (4%);
- целочисленное программирование (3%);
- теория очередей (систем массового обслуживания) (3%);
- прочие (6%).

2. Вероятностно-статистические аспекты имитационного моделирования

Часто математические модели интересующих нас систем содержат случайные параметры их функционирования и/или случайные внешние воздействия.

Пример 2.1. Исследуется работа парикмахерской.



Случайные внешние воздействия – интервалы времени между очередными посетителями $\delta_1, \delta_2, \dots$. Случайные параметры функционирования – времена $\tau_1^{(l)}$, $\tau_2^{(l)}$, ... обслуживания очередных своих клиентов l – м парикмахером. Цель исследования – решить, сколько нанять парикмахеров, чтобы среднее время ожидания в очереди не превысило заданной величины. Для этого нужно проиграть (ускоренно) несколько рабочих дней и посмотреть, какая будет очередь при разном числе парикмахеров. Чтобы проиграть день работы, потребуется сгенерировать конкретные значения d_1, d_2, \dots величин $\delta_1, \delta_2, \dots$. Будем считать, что $\delta_1, \delta_2, \dots$ – независимые, одинаково распределённые случайные величины: $F_{\delta_i}(d) \equiv F_{\delta}(d)$, т. е. d_1, d_2, \dots – реализация выборки из распределения $F_{\delta}(d)$.

Для некоторых случайных величин, скажем δ , есть основания применять определённый вид распределения (экспоненциальное $F_{\delta}(d) = 1 - e^{-\lambda d}$), для других ($\tau_1^{(l)}$) сам вид распределения неясен, и возникает задача проверки гипотезы о

том, что это некоторое фиксированное (стандартное) распределение, которое мы предполагаем.

2.1. Критерий согласия Колмогорова

Основная («нулевая») гипотеза состоит в том, что функция распределения случайной величины ξ есть некоторая фиксированная функция

$$H_0 : F_\xi(x) = F_0(x) \quad , \quad (1)$$

альтернативная гипотеза – функция распределения другая:

$$H_1 : F_\xi(x) \neq F_0(x) . \quad (2)$$

Для проверки этих гипотез используем выборку наблюдений величины ξ :

$$\{x_n\}, n = 1, \dots, N . \quad (3)$$

Эмпирическая функция распространения:

$$\hat{F}_N(x) = \frac{\text{число элементов выборки} \leq x}{N} . \text{ Если } H_0 \text{ справедлива, } \hat{F}_N(x) \text{ должна}$$

быть близка к $F_0(x)$ и наоборот, то есть показательной будет величина

$$D_N = \max_x \left| \hat{F}_N(x) - F_0(x) \right| - \text{максимальное отклонение эмпирической функции}$$

распределения от гипотетической. D_N случайна, так как выборка, по которой

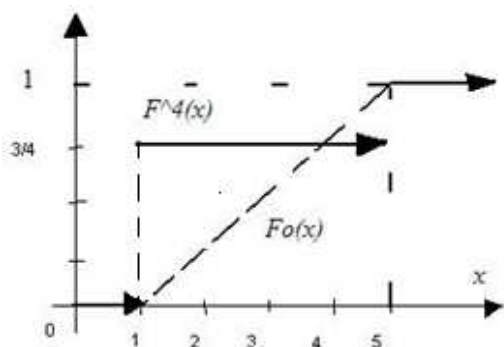
рассчитывается $\hat{F}_N(x)$, могла быть разной. Большие значения D_N в пользу H_1 ,

поэтому уровень значимости конкретных данных $\{x_n\}, n = 1, \dots, N$, которые привели к конкретному d_N , против гипотезы H_0 , равен

$$sl(d_N) = P\{D_N \geq d_N | H_0\} .$$

Теорема Колмогорова. Если $F_0(x)$ непрерывна, то распределение статистики D_N , при условии справедливости H_0 , не зависит от $F_0(x)$ и называется распределением Колмогорова с N степенями свободы.

Пример 2.2. По выборке $\{x_N\} = \{1, 1, 5, 1\}$ проверить гипотезу $H_0: F_0(x)$ – равномерное распределение на отрезке $[1, 5]$ ($U(1, 5)$).



Из рисунка видно, что $d_4 = 3/4$.
 $sl = P\{D_4 \geq 0.75 | H_0\}$.

По таблицам [7, таблица 6.2] находим $sl < 0.01$ (0.01 точно для 0.734). Такой уровень значимости данных интерпретируется как «высокая значимость», H_0 почти наверняка не подтверждается», ξ не имеет распределения $U(1, 5)$.

Теорема Колмогорова. $\lim_{N \rightarrow \infty} P\{\sqrt{N}D_N \leq x\} = K(x) = \sum_{l=-\infty}^{+\infty} (-1)^l e^{-2l^2 x^2}$.

$K(x)$ называется предельным распределением Колмогорова [7, таблица 6.1].

$P\{\sqrt{N}D_N \leq x\} \approx K(x)$ уже при $N > 20$.

Пример 2.2 (продолжение):

$$sl(d_N) = 1 - P\{\sqrt{N}D_N \leq \sqrt{N}d_N\} \approx 1 - K(\sqrt{N}d_N) = 1 - K(\sqrt{4} * 0.75) = 1 - K(1.5) = 1 - 0.9778 = 0.022.$$

Поскольку эмпирическая функция распределения является ступенчатой, а $F_0(x)$ неубывающей, то $\max |\hat{F}_N(x) - F_0(x)|$ приходится на одну из точек разрыва функции $\hat{F}_N(x)$.

2.2. Критерий согласия χ^2

Снова проверяются гипотезы (1), (2) по данным (3). Но теперь сравниваются плотность гипотетического распределения (H_0) и гистограмма (данные).

Алгоритм действий:

1. Разобьём область возможных значений случайной величины ξ по H_0 на r интервалов $[z_k, z_{k+1})$ $k = 1, \dots, r$.

2. Этим интервалам по распределению H_0 соответствуют вероятности попадания ξ в них $p_k = P\{z_k \leq \xi < z_{k+1}\}$.

3. v_k – число наблюдений (элементов выборки), попавших в k -й интервал

$$(v_1 + v_2 + \dots + v_r = N), \quad (4)$$

$\frac{v_k}{N}$ – частота попаданий в интервал.

4. Составим статистику $H^2 = \sum_{k=1}^r \frac{(v_k - Np_k)^2}{Np_k} = \sum_{k=1}^r \frac{N}{p_k} \left(\frac{v_k}{N} - p_k \right)^2$. При выпол-

нении H_0 частоты будут близки к соответствующим вероятностям, поэтому ожидаются малые значения статистики H^2 , поэтому большие значения свидетельствуют против H_0 .

5. $sl = P\{H^2 \geq h^2 | H_0\}$, h^2 – наблюдаемое (фактически полученное) значение статистики H^2 .

Теорема Пирсона. При справедливости H_0 распределение статистики H^2 стремится при $N \rightarrow \infty$ к распределению χ^2 с $r - 1$ степенями свободы.

Эвристическое доказательство:

по теореме Муавра – Лапласа $\frac{v_k - Np_k}{\sqrt{Np_k q_k}} \approx N(0,1)$,

$H^2 \approx (N(0,1))^2 + (N(0,1))^2 + \dots + (N(0,1))^2$, т. е. имеет распределение χ^2 с r степенями свободы (на самом деле $r - 1$, так как по (4) v_r зависит от остальных).

По теореме Пирсона

$$sl \approx P\{\chi_{r-1}^2 \geq h^2\} = 1 - P\{\chi_{r-1}^2 \leq h^2\} = 1 - \Phi_{r-1}^{\chi^2}(h^2). \quad (5)$$

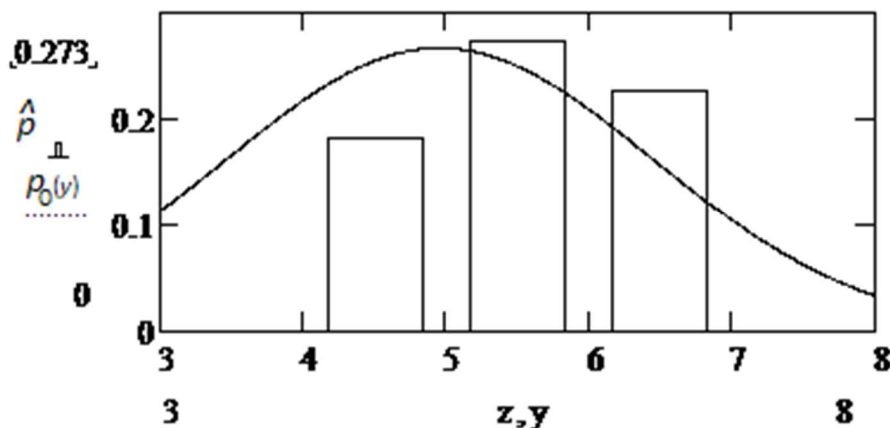
В силу приближенного характера формулы (5) желательно обеспечить $N \geq 50$, $n_k \geq 5 \div 10$, n_k – числа элементов выборки в интервалах (может быть, для этого придётся объединить интервалы).

Если $F_0(x)$ не полностью задана, т.е. l ее неизвестных параметров оценивались по исходной выборке, то число степеней свободы в (5) нужно уменьшить до $r - 1 - l$.

Пример 2.3. Выборка 3, 7, 5, 4, 8, 2, 5, 3, 4, 5, 4, 5, 6, 6, 3, 7, 5, 6, 5, 6, 4, 6. Согласуется ли она с гипотезой, что это наблюдения нормальной случайной величины?

Для визуализации выборки построим гистограмму. Вариационный ряд 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 7, 7, 8. Разобьём числовую ось на 5 интервалов $[-\infty, 4)$ $[4, 5)$ $[5, 6)$ $[6, 7)$ $[7, +\infty)$. Числа n_k попаданий в интервалы равны 4, 4, 6, 5, 3.

Определение. График $\hat{p} = n_k / N \Delta z_k$ как функции от \bar{z}_k – средних точек интервалов, называется гистограммой.



Проверим гипотезу H_0 , что $F_0(x)$ – нормальное распределение с параметрами, оценёнными по той же выборке $\hat{a} = 4.955$, $\hat{\sigma} = 1.495$ (плотность распределения $p_0(y)$).

$$p_1 = F_0(4) - F_0(-\infty) = 0.262, p_2 = F_0(5) - F_0(4) = 0.252, p_3 = 0.246, p_4 = 0.156, p_5 = 0.084.$$

$$h^2 = \frac{1}{N} \left(\sum_{k=1}^5 \frac{n_k^2}{p_k} - N^2 \right) = \frac{1}{22} \left(\frac{4^2}{0.262} + \dots + \frac{3^2}{0.084} - 22^2 \right) = 2.444,$$

$$sl \approx 1 - \Phi_{5-1-2}^{\chi^2}(2.444) = 1 - 0.705 = 0.295.$$

sl велик (> 0.1), данные не значимы против H_0 , это могли быть наблюдения нормальной случайной величины.

2.3. Генерация псевдослучайных чисел,

равномерно распределённых на отрезке $[0, 1]$

Для проведения имитации потребуется реализация

$$x_1, x_2, \dots, x_n, \dots, x_N \quad (6)$$

выборки $X_1, X_2, \dots, X_n, \dots, X_N$. Распределения нужны будут разные, для начала пусть это будет $U(0, 1)$.

Практически числа (6) получаем путём последовательных обращений к некоторому датчику, который может быть физическим устройством, связанным со случаем или компьютерной программой, которая вырабатывает разнообразные числа. В последнем случае эти числа называют псевдослучайными с распределением $F_\xi(x)$, если они удовлетворяют условиям, которые ожидаются от выборки из распределения $F_\xi(x)$. Псевдослучайными (якобы случайными), так как истинно случайными они не являются (могут быть всегда повторно получены).

Пример 2.4: 0.372, 0.893, 0.011, 3.62, 0.701 – не могут быть псевдослучайными числами с распределением $U(0, 1)$.

Пример 2.5: 0.1, 0.9, 0.7, 0.6, 0.3, 0.85, 0.69 – вряд ли годятся как псевдослучайные числа с распределением $U(0, 1)$, слишком много близких к единице. Это выявит тест Колмогорова.

Пример 2.6: 0.12, 0.97, 0.05, 0.81, 0.3, 0.75, 0.21, 0.59, 0.25, 0.63 – похоже, X_n не были независимыми (за большим следует малое и наоборот).

Таким образом, мы ожидаем (и нужно это проверять) от псевдослучайных чисел с распределением $U(0, 1)$:

- равномерность, т.е. согласие с распределением $U(0, 1)$.
- случайность.

2.4. Критерий случайности

Пусть есть упорядоченный набор чисел

$$x_1, \dots, x_N \quad (7)$$

(или *кортеж* – конечная последовательность).

Определение. Кортеж называется случайным, если он есть реализация выборки.

По данным (7) нужно проверить гипотезу $H_0 : X_1, \dots, X_N - i.i.d.$,

то есть независимые одинаково распределённые случайные величины, при альтернативе H_1 : это не так.

Рассмотрим вначале случай, когда элементы в (7) могут быть только двух видов (1 и 0, А и В, ...).

Пример 2.7. Кортеж 10100011010.

В действительно случайной выборке такого типа мы не ожидаем, что все «1» собьются вместе, и не ожидаем, что они будут регулярно чередоваться с «0».

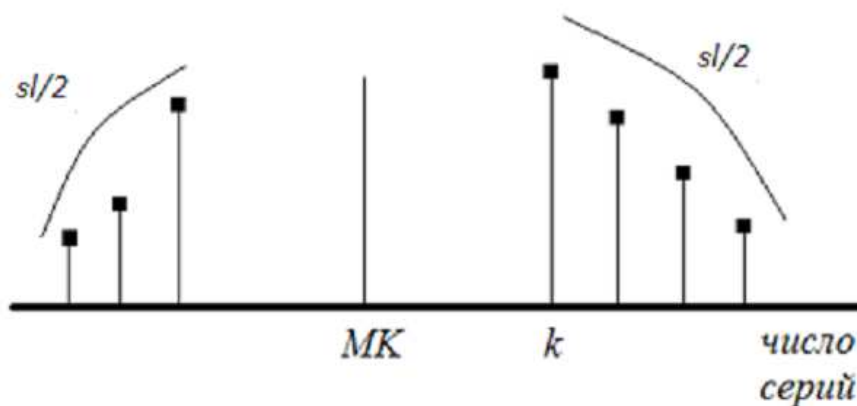
Серией в кортеже элементов двух видов называется совокупность идущих подряд одинаковых элементов, ограниченная противоположными элементами, или началом, или концом.

Статистикой критерия выберем случайное число серий K . Обозначим: N_1 – число элементов, которых в кортеже меньше; N_2 – число элементов, которых в кортеже больше.

Можно показать, что при справедливости H_0 $MK = 1 + 2 \frac{N_1 N_2}{N_1 + N_2}$. Значит,

сильные отклонения числа серий от MK – в пользу гипотезы H_1 .

$sl(k) = P\{K = k \text{ или ещё дальше от } MK \text{ (весь «хвост»)} + K \text{ принадлежит противоположному «хвосту» с такой же вероятностью} | H_0\}$.



При фиксированном $sl = \alpha$ можно табулировать величины $k_{\min}^{(\alpha)}$ и $k_{\max}^{(\alpha)}$ такие, что при $k_{\min}^{(\alpha)} < k < k_{\max}^{(\alpha)}$ достигаемый sl будет $> \alpha$.

Пример 2.7 (продолжение): $N_1 = 5, N_2 = 6$, число серий $k = 8$.

По таблице распределения K [7, таблица 6.7]:

$$k_{\min}^{(0.05)} = 3 < 8 < 10 = k_{\max}^{(0.05)} \Rightarrow sl > 0.05;$$

$$k_{\min}^{(0.1)} = 3 < 8 < 10 = k_{\max}^{(0.1)} \Rightarrow sl > 0.1.$$

Значит, $sl > 0.1$, данные не значимы против H_0 , кортеж случаен.

Если либо N_1 , либо $N_2 > 20$, то статистика

$$Z = \frac{(K - MK) - \frac{1}{2} \text{sign}(K - MK)}{\sqrt{\frac{2N_1N_2(2N_1N_2 - (N_1 + N_2))}{(N_1 + N_2)^2(N_1 + N_2 - 1)}}}$$

при H_0 имеет приближённо нормальное стандартное распределение. Поэтому

$$sl(z) \approx 2(1 - \Phi(|z|)). \quad (8)$$

Теперь пусть числа (7) – якобы выборка из произвольного непрерывного распределения. Решить вопрос о не случайности этого кортежа можно следующим образом.

Вспомним: медиана – квантиль порядка 0.5; выборочная медиана \widehat{Me} равна срединному элементу вариационного ряда, если число наблюдений нечётно, и равна полу сумме 2-х срединных элементов вариационного ряда, если число наблюдений чётно. Составим разности $(x_n - \widehat{Me}) \Rightarrow \begin{cases} +, \geq 0 \\ -, < 0 \end{cases}$, тогда вместо (7) имеем кортеж знаков $+$ и $-$, случайность которого умеем проверять.

Если кортеж (7) случайный, то последовательные превышения медианы есть независимые события и кортеж знаков будет случайным, поэтому если кортеж знаков неслучайный, то и изначальный кортеж не был случайным.

2.5. Мультипликативный конгруэнтный метод генерации псевдослучайных чисел

Пусть x_{n-1} – некоторое число из $(0, 1)$. Получим $x_n = \mathcal{D}(\beta x_{n-1})$, где β – большое целое число; \mathcal{D} – взятие дробной части (откинуть целую часть, если она есть). Новое число x_n опять из $(0,1)$, его опять умножить на β и т. д., получится последовательность чисел, равномерно распределённых на $(0, 1)$.

Алгоритм можно переписать. Пусть a_n – целое, M – большое целое,

$$x_n = \frac{a_n}{M}, \quad (9)$$

$$a_n = \mathcal{D}\left(\frac{\beta a_{n-1}}{M}\right)M. \quad (10)$$

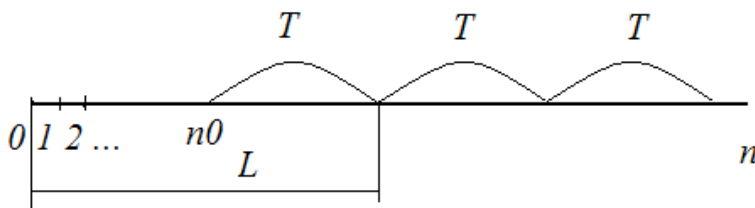
Определение. Остаток от деления натурального числа p на натуральное число q обозначается $(p) \bmod q$.

Формулу (10) перепишем

$$a_n = (\beta a_{n-1}) \bmod M, \quad n = 1, 2, \dots \quad (11)$$

Алгоритм (11), (9) легче реализуется на компьютере.

Последовательность (11) всегда зацикливается, т. е. начиная с некоторого номера n_0 образуется период длины T , который затем бесконечно повторяется. L – длина отрезка аперiodичности.



Пример 2.8. При $a_0 = 65539$, $\beta = 65539$, $M = 2^{31}$ из теоретико-числовых соображений следует: $T = M/4 = 536870912$.

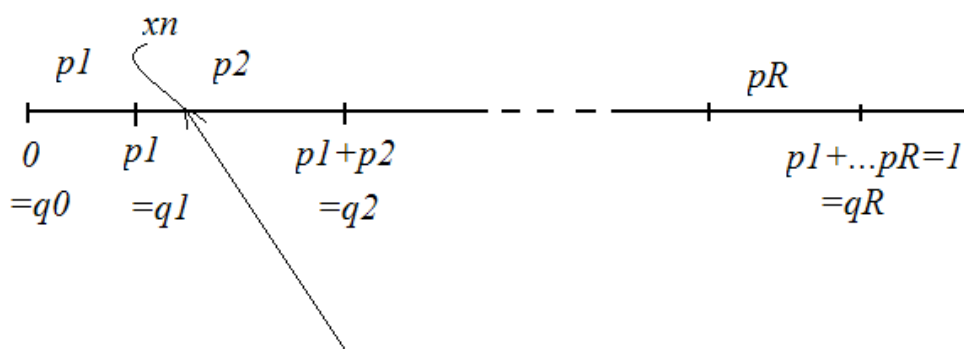
2.6. Моделирование дискретной случайной величины

Пусть нужно получить реализацию y_1, \dots, y_N выборки из η :

$y^{(r)}$	$y^{(1)}$	$y^{(2)}$	\dots	$y^{(R)}$
p_r	p_1	p_2	\dots	p_R

– закон распределения.

Изобразим вероятности из этого закона на числовой оси.



Далее идёт равномерная «стрельба» по отрезку $[0, 1]$:

- 1) $n = 1$;
- 2) генерируется число x_n – реализация $\xi \sim U(0,1)$;
- 3) объявляется реализация величины η $y_n = y^{(r)}, r: q_{r-1} < x_n \leq q_r$;
- 4) $n = n + 1$, переход к пункту 2.

Поскольку ξ равномерно распределена на $[0, 1]$, значения $y^{(r)}$ будут появляться с вероятностями, равными длине отрезка $[q_{r-1}, q_r]$, т. е. p_r .

2.7. Моделирование непрерывной случайной величины с произвольным распределением

Пусть нужно получить реализацию y_1, \dots, y_N выборки из η : $F_\eta(y) = F_0(y)$ – заданная строго монотонно возрастающая функция.

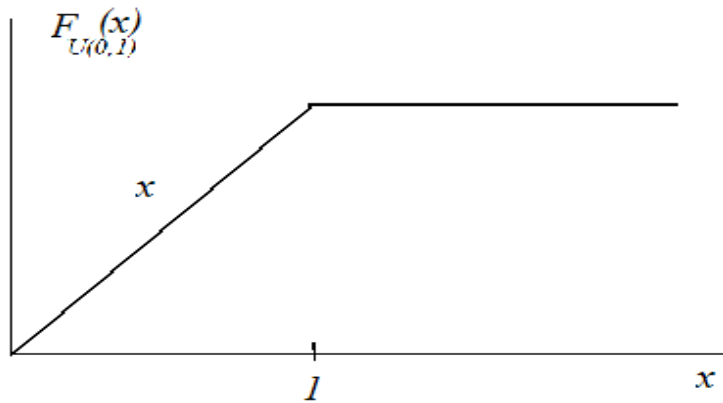
Теорема (метод обратной функции). Пусть $x = F_0(y)$, $y = F_0^{-1}(x)$ – обратная функция. Если $\xi \sim U(0,1)$, то случайная величина $\eta = F_0^{-1}(\xi)$ будет иметь (нужное) распределение $F_0(y)$.

Доказательство: $F_\eta(y) = P\{\eta \leq y\} = P\{F_0^{-1}(\xi) \leq y\} =$

$= P\{F_0(F_0^{-1}(\xi)) \leq F_0(y)\} = P\{\xi \leq F_0(y)\} = F_{U(0,1)}[F_0(y)] = F_0(y)$. Третье равенство

имеет место, так как при строго монотонно возрастающем преобразовании

знак неравенства сохраняется. Последнее равенство очевидно из рисунка.



Пример 2.9. Смоделировать η с $M(\lambda)$ – показательное распределение с параметром λ , то есть $x = F_0(y) = 1 - e^{-\lambda y}$. Найдём обратную функцию $e^{-\lambda y} = 1 - x$, $-\lambda y = \ln(1 - x)$, $y = -\frac{1}{\lambda} \ln(1 - x) = F_0^{-1}(x)$.

Поэтому $y_n = -\frac{1}{\lambda} \ln(1 - x_n)$, где x_n – псевдослучайные числа с $U(0, 1)$, есть реализации η с $M(\lambda)$.

Есть и другие методы моделирования случайной величины с заданным непрерывным распределением (для моделирования нормальной случайной величины метод обратной функции непригоден аналитически, но в *MathCAD* есть обратная функция *qnorm()*).

2.8. Моделирование случайных векторов

Пусть требуется сгенерировать выборку $\vec{y}^1, \vec{y}^2, \dots, \vec{y}^N$ N значений случайного

вектора $\vec{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_r \end{pmatrix}$ с заданными характеристиками:

математическим ожиданием $M \vec{Y} = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_r \end{pmatrix}$ и ковариационной матрицей

$$K_{\vec{Y}} = M \left(\vec{Y} - M \vec{Y} \right) \left(\vec{Y} - M \vec{Y} \right)^T = \begin{pmatrix} K_{11} & K_{12} & \dots & K_{1r} \\ K_{21} & K_{22} & \dots & K_{2r} \\ \dots & \dots & \dots & \dots \end{pmatrix}.$$

Построим вектор \vec{Y} с помощью вектора $\vec{X} = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_r \end{pmatrix}$ с независимыми, одинаково

распределёнными компонентами с нулевым математическим ожиданием и единичной дисперсией. Например, при $r = 2$ запишем:

$$\begin{aligned} Y_1 &= c_{11}X_1 + a_1; \\ Y_2 &= c_{21}X_1 + c_{22}X_2 + a_2. \end{aligned} \quad (12)$$

Очевидно, Y_1 и Y_2 имеют нужные математические ожидания. Если $c_{22} = 0$, то они функционально (полностью) зависимы, если $c_{21} = 0$, то они независимы.

Подберём c_{ij} так, чтобы Y_1 и Y_2 имели нужные стандартные отклонения и коэффициент корреляции.

Пример 2.10. Требуется сгенерировать выборку значений двумерного нормального случайного вектора с математическими ожиданиями компонент 2 и 5, стандартными отклонениями 1 и 4 и коэффициентом корреляции компонент 0.5.

Решение в *MathCAD*.

$$\begin{aligned} \text{ORIGIN} &:= 1 \\ N &:= 1000 \quad a := \begin{pmatrix} 2 \\ 5 \end{pmatrix} \quad \sigma_1 := 1 \quad \sigma_2 := 4 \quad \rho := 0.5 \end{aligned}$$

Когда случайные величины были скалярными, выборки их располагались в *MathCAD* в столбец, т. е. вектор. Теперь сами случайные величины – векторы, т. е. столбцы. Поэтому выборки их будем располагать в строку, т. е. каждое новое

наблюдение – новый столбец в матрице размера $r * N$, где r – размерность случайного вектора.

Генерируем выборку вектора \vec{X} :

$n := 1..N$

$$x^{(n)} := \begin{pmatrix} \text{rnorm}(1, 0, 1)_1 \\ \text{rnorm}(1, 0, 1)_1 \end{pmatrix}$$

Первые два элемента выборки:

$$x^{(1)} = \begin{pmatrix} 1.107 \\ 0.662 \end{pmatrix} \quad x^{(2)} = \begin{pmatrix} -2.039 \\ 0.964 \end{pmatrix} .$$

Вся выборка:

$$X =$$

	1	2	3	4	5	6
1	1.107	-2.039	1.697	-0.856	0.635	-0.728
2	0.662	0.964	-0.667	0.045	-1.611	...

По формулам (12) получаем требуемую выборку вектора \vec{Y} .

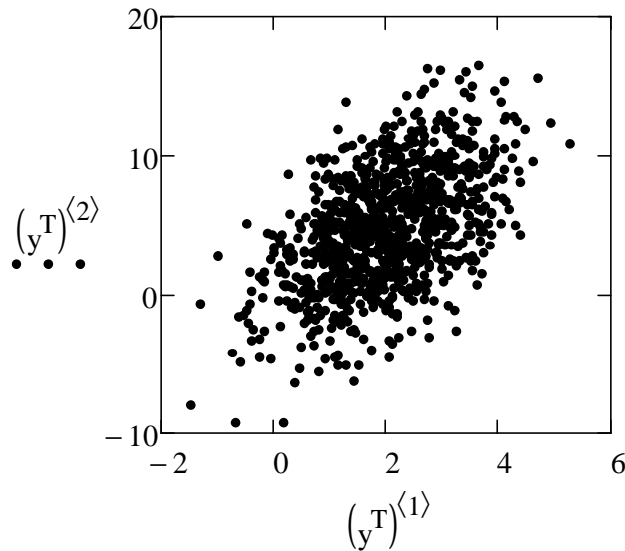
$$y^{(n)} := \begin{bmatrix} (x^{(n)})_1 \cdot \sigma_1 + a_1 \\ (x^{(n)})_1 \cdot \sigma_2 \cdot \rho + (x^{(n)})_2 \cdot \sigma_2 \cdot \sqrt{1 - \rho^2} + a_2 \end{bmatrix}, \quad (13)$$

$$Y =$$

	1	2	3	4	5	6
1	3.107	-0.039	3.697	1.144	2.635	1.272
2	9.508	4.26	6.082	3.446	0.69	...

Для проверки оценим по выборке коэффициент корреляции компонент вектора и построим корреляционное поле (диаграмму рассеивания):

$$\text{corr}[(y^T)^{(1)}, (y^T)^{(2)}] = 0.528 .$$



Или другой метод генерации коррелированных нормальных величин.

Пусть вектор $\vec{X} \sim N(\vec{0}, I)$. Тогда

$$\vec{Y} = A\vec{X} + \vec{a} \sim N(\vec{a}, K), \quad (13.1)$$

где $A = S\Lambda^{1/2}$. Λ – диагональная матрица из собственных значений, S – матрица из ортонормированных собственных векторов требуемой ковариационной матрицы K .

2.9. Статистический анализ результатов моделирования

После того как мы позаботились о случайности и правильной распределённости случайных внешних воздействий и/или параметров модели, наша модель будет «жить» почти реальной жизнью и фиксировать её разнообразные величины от прогона к прогону – это все равно что наблюдать многократно реальную систему. А значит, к этим зафиксированным данным можно и должно применять весь арсенал математической статистики для оценок характеристик интересующих нас величин, их зависимостей (корреляции) и т. д.

2.10. Упражнения и лабораторный практикум

Упражнение 2.1.

Выполнить упражнение 9.25 из [21] по вариантам. Повторить решение в *Statistica*. Например, для 13-го варианта: *Statistics, Distrib. Fitting, Exponent, Var V1, Options, Kolmogorov – Smirnov, Yes (continu...), Graph, Cumulative, Relative (%)*, *Parameters, Low = 121, Upper = 144, Number = 30, Plot of observ...*

Упражнение 2.2.

Найти значение статистики Колмогорова для выборки $\{1, 1, 5, 1\}$ и гипотетического распределения $N(1, 1)$.

Упражнение 2.2а

Найти значение статистики Колмогорова для выборки $\{2, 1, 5, 2\}$ и гипотетического распределения $U(0, 5)$. Ответ 0.350.

Упражнение 2.3.

В примере 2.3 подтвердить вычисления вероятностей p_k , формулу для h^2 .

Упражнение 2.4.

Выполнить упражнение 9.24 из [21] по вариантам.

Упражнение 2.5.

При прослушивании Космоса в поисках посланий внеземных цивилизаций была получена последовательность «точек» и «тире»: (-.-.-.-.-.....-.-.-.-.-....-.). Можно ли считать, что последовательность не случайна, а несет закодированное послание? Использовать точное распределение, а затем нормальную аппроксимацию.

Упражнение 2.6.

Проверить случайность кортежа из примера 2.6.

Упражнение 2.6a

Проверка гипотезы случайности кортежа 1, 4, 3, 7, 5, 4. Каков уровень значимости, рассчитанный по приближённой формуле для больших кортежей? Ответ 0.860.

Упражнение 2.7.

Выполнить упражнение 9.27 из [21] по вариантам.

Упражнение 2.8.

При множителе при seed равном $b = 21$ и $seed = 11$ (соответствует первому члену последовательности), делителе равном $M = 111$, найти мультипликативным конгруэнтным методом третье псевдослучайное число с равномерным распределением на отрезке $[0, 1]$. Ответ 0.703.

Упражнение 2.9.

Используя мультипликативный конгруэнтный генератор с параметрами $\beta = 125$, $M = 8192$, $seed = 1$, получить 5 первых псевдослучайных чисел.

Упражнение 2.10.

Используя результаты упражнения 2.9, смоделировать подбрасывание игровой кости.

Упражнение 2.11.

Используя результаты упражнения 2.9, сгенерировать псевдослучайные числа $\sim M(1)$.

Упражнение 2.11а

При моделировании подбрасывания игральной кости, очередное псевдослучайное число из $U(0, 1)$ оказалось $= 0.837$. Какой гранью выпадет кость?

Ответ 6.

Упражнение 2.11б

При получении пс. ч. из показательного распределения с параметром 1.3 методом обратной функции используется пс. ч. из равномерного распределения на $(0, 1)$ равное 0.976. Каков будет результат? Ответ 2.87.

Упражнение 2.11в

При получении пс. ч. из равномерного распределения на $(100, 300)$ методом обратной функции используется пс. ч. из равномерного распределения на $(0, 1)$ равное 0.976. Каков будет результат? Ответ 295.

Упражнение 2.11г

При получении пс. ч. из пуассоновского распределения с параметром 6 методом обратной функции используется пс. ч. из равномерного распределения на $(0, 1)$ равное 0.131. Каков будет результат? Ответ 3.

Упражнение 2.12.

Записать формулу получения псевдослучайных чисел, распределённых на отрезке $[0, 2]$ с плотностью $\sim \sqrt{x}$, методом обратной функции.

Упражнение 2.12а.

Найти методом обратной функции псевдослучайное число, из распределения на отрезке $[0, 1]$ с плотностью, пропорциональной корню квадратному из аргумента. Исходное псевдослучайное число из $U(0, 1)$ равно 0.123. Ответ 0.247.

Упражнение 2.13.

Доказать формулу (8) для $sl(z)$ критерия случайности.

Упражнение 2.14.

Получить формулы (13).

Упражнение 2.15.

Доказать (13.1).

3. Пакет *MathCAD*

MathCAD – один из математических пакетов общего назначения (есть еще *Matlab*, *Mathematica* и др.). *CAD* (*Computer-Aided Design*) – проектирование с помощью компьютера. Основное преимущество этого пакета в том, что математическая запись в *MathCAD* максимально близка к общематематической, только страницы получаются «живые».

Пример 3.1.

Набрали $\int_0^1 x^2 dx = 0.333$, поставили знак $=$, и получился ответ.

Дальнейшее знакомство с пакетом будет на примере.

Пример: критерий согласия χ^2 и гистограмма

В начале чистого рабочего листа (*worksheet*), с клавиатуры набираем *ORIGIN*: = 1. *ORIGIN* (источник, начало) нумерации элементов массивов будет установлено в 1 (по умолчанию нумерация начинается с 0). Оператор присваивания «:=» берется или из арифметической палитры (*Calculator*), или вводится с клавиатуры нажатием «:=».

В вектор x вводим выборку. В матричной палитре выбираем *Insert matrix*, задаем 1 строку и 22 столбца. Транспонирование – там же.

$x := (3 \ 7 \ 5 \ 4 \ 8 \ 2 \ 5 \ 3 \ 4 \ 5 \ 4 \ 5 \ 6 \ 6 \ 3 \ 7 \ 5 \ 6 \ 5 \ 6 \ 4 \ 6)^T$

Проверка:

$N := \text{length}(x), \quad N = 22.$

$r := 5, \quad \text{int} := (-10 \ 4 \ 5 \ 6 \ 7 \ 20)^T, \quad \text{hist}(\text{int}, x)^T = (4 \ 4 \ 6 \ 5 \ 3).$

$\text{hist}(\text{int}, x)$ – функция из палитры функций $f(x)$ возвращает вектор чисел элементов из x , попавших в интервалы с границами, указанными в векторе int .

Задаем цикл. Символ «...» берётся или из матричной палитры или вводится с клавиатуры нажатием «;».

$$i := 1..r, \quad \text{delt}_i := \text{int}_{i+1} - \text{int}_i, \quad f_i := \frac{\text{hist}(\text{int}, x)_i}{N \cdot \text{delt}_i}, \quad z_i := \frac{\text{int}_i + \text{int}_{i+1}}{2}.$$

Строим график гистограммы. *Insert*→*Graph*→*X-Yplot*. По горизонтали откладываем z , по вертикали f . *Format*→*Traces*→*Type*→*Bar*. Получим график из раздела 2.2.

$$k := 1..r, \quad p_k := \text{pnorm}(\text{int}_{k+1}, 4.95, 1.49) - \text{pnorm}(\text{int}_k, 4.95, 1.49).$$

$\text{pnorm}(x, a, \sigma)$ – функция нормального распределения с параметрами a, σ на аргументе x .

$$p^T = (0.262 \quad 0.252 \quad 0.246 \quad 0.156 \quad 0.084) .$$

$$\text{hsq} := \frac{1}{N} \left[\sum_{k=1}^r \frac{(\text{hist}(\text{int}, x)_k)^2}{p_k} \right] - N, \quad \text{hsq} = 2.444,$$

$$\text{SL} := 1 - \text{pchisq}(\text{hsq}, r - 1 - 2), \quad \text{sl} = 0.295.$$

3.1. Упражнения и лабораторный практикум

Упражнение 3.1.

Повторить в *MathCAD* пример 2.3.

Упражнение 3.2.

Методом обратной функции из псевдослучайных чисел (пс.ч.) $\sim U(0, 1)$ (использовать функцию $\text{runif}(n, 0, 1)$), получить пс.ч. $\sim U(5, 20)$ и пс.ч. $\sim M(2)$. Проиллюстрировать результаты гистограммами (с использованием функции $\text{histogram}(x, y)$). Наблюдая гистограммы, несколько раз пересчитать рабочий лист (*Tools, Calculate, Calculate Worksheet*). Что изменилось? Почему? Что сохранилось? Найти оценки математических ожиданий и стандартных отклонений. Сравнить их с истинными значениями, убедиться в несмещённости и состоятель-

ности. Найти в *MathCAD* встроенные статистические функции оценки математического ожидания и стандартного отклонения. Вместо явно найденных обратных функций (квантилей) применить их эквиваленты в *MathCAD* *qunif()* и *qexp()*.

Получить пс.ч. $\sim N(1, 3)$ (с использованием функции *rnorm(..., ..., ...)*). Построить гистограмму. Найти оценки математического ожидания и стандартного отклонения. Сравнить их с истинными значениями.

Упражнение 3.2.а

Пусть $Y = U_1 + U_2 + \dots + U_{12} - 6$, где $U_i - i.i.d. \sim U(0,1)$. Найти её м.о. и дисперсию. Найти то же для похожей величины $Y_d = 12 * U(0,1) - 6$.

По ЦПТ Y почти нормален (и стандартен). Проверить гистограммой, сгенерировав выборку. Оценить м.о. и дисперсию. Почему 12 слагаемых, казалось бы, чем больше - тем лучше. Если взять 120 слагаемых и вычесть(??) Что изменится?

Упражнение 3.3.

Проверить, что стандартное отклонение выборочного среднего в должное число раз меньше такового для самой величины. Для этого сгенерировать N значений величины и N значений выборочного среднего (на основе M набл.) и построить их гистограммы; вначале использовать нормальное распределение, затем равномерное. Объяснить результаты с помощью центральной предельной теоремы.

Упражнение 3.4.

Методом «стрельбы» по отрезку $(0, 1)$ получить выборку из распределения *Hurgeo*(M, N, p), $M = 10, N = 3, p = 0.4$. (Такое распределение имеет случайная величина M_N – число дефектных изделий в контрольной выборке (бесповторной) N изделий из партии объёмом M с долей дефектных изделий p .

$$P\{M_N = m\} = \begin{cases} \frac{C_{Mp}^m C_{M-Mp}^{N-m}}{C_M^N}, \max\{0, Mp + N - M\} \leq m \leq \min\{Mp, N\} \\ 0, \text{ в других случаях} \end{cases} . MM_N = Np,$$

$DM_N = Np(1-p) \frac{M-N}{M-1}$). Построить гистограмму. Повторить, используя функцию $qhypergeom(x, Mp, M - Mp, N)$.

Упражнение 3.4а.

Методом обратной функции с помощью $qpois()$ сгенерировать пуассоновские п.с.ч. с параметром 3. Построить гистограмму. Найти оценки м. о. и стандартного отклонения. Сравнить с истинными значениями.

Упражнение 3.5.

Сгенерировать 1000 наблюдений двумерного нормального случайного вектора. Математические ожидания компонент 2 и 5, стандартные отклонения 1 и 4, коэффициент корреляции 0.5 (какова требуемая ковариационная матрица?). Построить диаграмму рассеивания, оценить коэффициент корреляции, как для исходных, так и для результирующих величин. Увеличить объём выборки. Как ведёт себя оценка коэффициента корреляции? Как называется это свойство оценки? Изменить коэффициент корреляции на (-0.9) , 1.

Упражнение 3.6.

Используя $gnorm(N,0,1)$ сгенерировать выборку 3-мерного вектора \vec{Y} с независимыми нормальными стандартными компонентами. Убедиться, что $\vec{Z} = A \cdot \vec{Y} + \vec{\mu}$ где $A =$

-0.224	0.951	-0.212
-1.13	0.828	0.193
2.171	0.529	0.079

будет иметь распределение $N(\vec{\mu}, K)$ где $K =$

1	1	0
1	2	-2
0	-2	5

Упражнение 3.7.

Для моделирования 2-х коррелированных пуассоновских величин:

1. сгенерировать методом п. 2.8 выборку 2-мерного вектора \vec{Z} с нормальными стандартными компонентами и заданной корреляционной матрицей.
2. для каждой компоненты применить суперпозицию преобразований $W_i = \text{qpois}(\text{pnorm}(Z_i, 0, 1), a_i)$ где a_i – требуемый параметр пуассоновской компоненты. В сущности это метод обратной функции с использованием должным образом коррелированных равномерно распределённых в $(0, 1)$ исходных п.с.ч.
3. оценить параметры полученного вектора, убедиться, что они должные.

Для моделирования произвольного числа коррелированных пуассоновских величин можно воспользоваться методом *Упр. 3.6.*

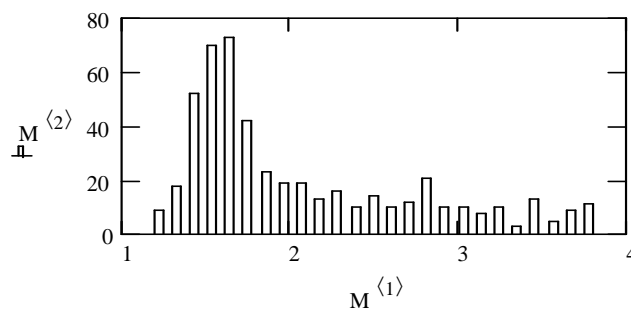
4. Примеры имитационного моделирования

4.1. Простая система с влиянием внешней среды

Пример 4.1. Известно, что на систему воздействует входная случайная величина $L \sim U(-1, 2)$ и влияет внешняя среда, случайная величина $F \sim N(2, 1)$. L и F независимы. Нас интересует выходная величина $Y = \sqrt{(1 + e^{-L})^2 + (1 - e^{-F})^2}$.

Что можно сказать о Y ? Ясно, что это случайная величина, распределение которой дедуктивно найти очень сложно. Прежде всего, нас интересует математическое ожидание MY . Для оценки математического ожидания сгенерируем выборку y_1, \dots, y_N значений величины Y . Затем $\widehat{MY} = \bar{y} = \frac{1}{N} \sum_{n=1}^N y_n$. Алгоритм: генерация выборки $l = (l_1, \dots, l_N)^T$; генерация выборки $f = (f_1, \dots, f_N)^T$; для $n = 1, \dots, N$, $y_n = \sqrt{(1 + e^{-l_n})^2 + (1 - e^{-f_n})^2}$.

Если провести моделирование при $N = 500$, то получим оценку математического ожидания $\widehat{MY} = 2.07$, оценку стандартного отклонения $\widehat{\sigma}_Y = 0.687$. Гистограмма даёт представление о распределении выходной величины Y . В частности, мода вблизи 1.5 и т. д.



4.2. Доверительная оценка погрешности имитационных расчётов

Если целью является нахождение математического ожидания MY , то можно указать погрешность его определения. По центральной предельной теореме при больших N , независимо от вида распределения Y , \bar{Y} будет иметь приближённо нормальное распределение: $\bar{Y} \approx N(M\bar{Y}, \sigma_{\bar{Y}}) = N(MY, \frac{\sigma_Y}{\sqrt{N}})$. Значит, \bar{y} не может случиться дальше от MY , чем на $(2 \div 3)$ стандартных отклонения:

$\bar{y} = MY \pm (2 \div 3) \frac{\sigma_Y}{\sqrt{N}}$, что эквивалентно

$$MY = \bar{y} \pm (2 \div 3) \frac{\sigma_Y}{\sqrt{N}}. \quad (14)$$

Здесь множитель 2 соответствует доверительной вероятности 0.954, а 3 – 0.997.

Пример 4.1(продолжение). Заменяя стандартное отклонение его оценкой $\hat{\sigma}_Y = 0.687$, получим: $MY = 2.07 \pm 0.06$ с вероятностью 0.954.

По (14) можно оценить необходимое число прогонов N для достижения заданной точности.

4.3. Вычисление определённых и кратных интегралов методом Монте-Карло

Пусть нужно вычислить определённый интеграл $I = \int_a^b f(x)dx$. Вспомним теорему о среднем.

$\exists c : I = (b-a)f(c)$. Точка c неизвестна. Что, если взять её случайным образом, равномерно распределённой на отрезке интегрирования?

$c \rightarrow \xi \sim U(a, b)$. Тогда и результат формулы теоремы о среднем станет случайным: $\eta = (b-a)f(\xi)$. Каково его математическое ожидание?

$$M\eta = M(b-a)f(\xi) = (b-a)Mf(\xi) =$$

$$= (b-a) \int_{-\infty}^{+\infty} f(x)p_{\xi}(x)dx = (b-a) \int_a^b f(x) \frac{1}{b-a} dx = \int_a^b f(x)dx = I, \text{ то есть совпадает}$$

с искомым интегралом!

Мы знаем, что математическое ожидание случайной величины можно оценить с контролируемой точностью по её сгенерированной выборке (14), поэтому

$$y_n = (b-a)f(x_n), \quad x_n - \text{н.с.ч.} \sim U(a, b); \quad (15)$$

$$I = \frac{1}{N} \sum_{n=1}^N y_n \pm (2 \div 3) \frac{\sigma_{\eta}}{\sqrt{N}}, \quad (16)$$

σ_{η} можно грубо оценить сверху:

$$\sigma_{\eta}^2 = D\eta = D(b-a)f(\xi) =$$

$$= (b-a)^2 Df(\xi) = (b-a)^2 (Mf^2(\xi) - (Mf(\xi))^2) \leq$$

$$\leq (b-a)^2 Mf^2(\xi) = (b-a)^2 \int_{-\infty}^{+\infty} f^2(x)p_{\xi}(x)dx =$$

$$= (b-a)^2 \int_a^b f^2(x) \frac{1}{b-a} dx \leq \|f(x)\| \leq d, \text{ на } [a, b] \leq (b-a)^2 d^2.$$

$$\sigma_{\eta} \leq (b-a)d. \quad (16.1)$$

Или заменить оценкой стандартного отклонения по выборке (15).

Пример 4.2. Вычислить интеграл $I = \int_0^{\pi/2} \sqrt{\sin(x)} dx$.

Используем числа t_n , равномерно распределённые на $[0, 1]$ из таблицы случайных чисел. Из них получаем $x_n \sim U(0, \pi/2)$.

n	1	2	3	4	5	6	7	8	9	10
t_n	0.1	0.09	0.73	0.25	0.33	0.76	0.52	0.01	0.35	0.86
x_n	0.157	0.141	1.147	0.393	0.518	1.194	0.817	0.016	0.55	1.351
y_n	0.621	0.59	1.5	0.972	1.106	1.515	1.341	0.197	1.135	1.552

По формуле (16) $I = 1.053 \pm 2 \frac{\sigma_\eta}{\sqrt{10}}$ с доверительной вероятностью 0.954. По

(16.1) $\sigma_\eta \leq 1.57$. В этом примере можно дать более точную оценку: $\sigma_\eta \leq 1.253$.

Таким образом $I = 1.05 \pm 0.79$.

Оценка по выборке $\hat{\sigma}_\eta = 0.46$, тогда $I = 1.05 \pm 0.29$.

Иногда в качестве погрешности берут вероятную ошибку $0.674 \frac{\sigma_\eta}{\sqrt{N}}$, т. е.

0.674 вместо $(2 \div 3)$. В чём смысл вероятной ошибки?

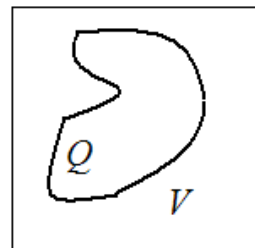
0.674σ	?	2σ	?
?	0.683	0.954	?

В том, что вероятности того, что фактическая ошибка будет по абсолютной величине больше этой ошибки или меньше, одинаковы ($= 0.5$). Тогда $I = 1.053 \pm 0.098$.

Если нужно вычислить кратный интеграл $I = \int_V f(\vec{x}) d\vec{x}$, где $V - s$ – мерный

куб, то расчет идёт по формуле (16), где $y_n = |V| f(\vec{x}^n)$, \vec{x}^n – выборка из $\vec{\xi}$ – равномерно распределённого в кубе V случайного вектора; σ_η оценивается аналогично.

Для вычисления $I = \int_Q f(\vec{x}) d\vec{x}$, где Q :



введём $\tilde{f}(\vec{x}) = \begin{cases} f(\vec{x}), \vec{x} \in Q \\ 0, \vec{x} \notin Q \end{cases}$, тогда $I = \int_V \tilde{f}(\vec{x}) d\vec{x}$.

Метод Монте-Карло экономнее в вычислениях, чем квадратурные формулы в случае кратных интегралов, и проще учитывает сложную форму области интегрирования Q .

4.4. Имитационное моделирование инвестиционных рисков

Пример 4.3. Фирма рассматривает инвестиционный проект по производству продукта «А». Проект рассчитан на $T = 5$ лет.

Первоначальные инвестиции (в момент $t = 0$) $I_0 = 4200$ условных денежных единиц, затем в моменты $t = 1, \dots, T$ будут поступления (чистые денежные потоки) NCF_t . Критерием качества инвестиционного проекта часто берут чистую современную ценность

$$NPV = \sum_{t=1}^T \frac{NCF_t}{(1+r)^t} - I_0.$$

Дисконтный множитель $\frac{1}{1+r}$ пересчитывает величину поступления X_1 через

год, к $X_0 = \frac{1}{1+r} X_1$, т. е. к её нынешней эквивалентной величине. Действительно,

положив X_0 в банк со ставкой r , через год как раз будет величина $X_1 = (1+r)X_0$

. Через 2 года $X_2 = (1+r)X_1 = (1+r)^2 X_0$, т. е. величина X_2 эквивалентна нынеш-

ней $X_0 = \frac{1}{(1+r)^2} X_2$ и т. д.; r – норма дисконта (перерасчётов). Ясно, что если

$NPV < 0$, то следует отказаться от проекта (он убыточен).

Рассмотрим структуру поступлений NCF_t .

$$NCF_t = [Q_t(P_t - V_t) - F - A](1 - Tax) + A,$$

где Q_t – объем выпуска (шт.) – будем считать случайной величиной $\sim U(150, 300)$;

P_t – цена за штуку $\sim U(40, 55)$;

V_t – переменные затраты на штуку $\sim U(25, 35)$;

F – постоянные затраты ($= 500$), A – амортизация ($= 100$), Tax – налог на прибыль ($= 0.6$, или 60%).

Поскольку в NCF_t входят случайные факторы, то и NPV будет случаен. Хотелось бы знать его математическое ожидание $M(NPV)$ и стандартное отклонение σ_{NPV} .

Риск можно измерить коэффициентом вариации $CV = \sigma_{NPV}/M(NPV)$. Ясно, что чем $M(NPV)$ больше и σ_{NPV} меньше, тем лучше.

Другая мера риска инвестиционного проекта $R = P\{NPV < 0\}$.

Алгоритм имитаций:

- 1) $n = 1$;
- 2) генерация Q – вектора-выборки T объёмов выпуска (в каждом из T лет);
- 3) генерация P – вектора-выборки T возможных цен;
- 4) генерация V – вектора-выборки T возможных переменных затрат;
- 5) подсчет $NPV_n = \sum_{t=1}^T \frac{[Q_t(P_t - V_t) - F - A](1 - Tax) + A}{(1 + r)^t} - I_0$ – чистая современная ценность проекта в n -м прогоне;
- 6) $n = N$? Если нет – $n = n + 1$, переход к п. 2, если да – переход к п. 7;
- 7) *Stop*, готов NPV – вектор-выборка N результатов имитации. Статистический анализ результатов.

Решение в MathCAD

ORIGIN := 1 T := 5 I0 := 4200 r := 0.1 F := 500 A := 100 Tax := 0.6

N := 1000 – число имитаций

<u>NPV</u> :=	<div style="margin-left: 20px;">for $n \in 1..N$</div> <div style="margin-left: 40px;"> $Q \leftarrow \text{runif}(T, 150, 300)$ $P \leftarrow \text{runif}(T, 40, 55)$ $V \leftarrow \text{runif}(T, 25, 35)$ </div> <div style="margin-left: 40px;"> $NPV_n \leftarrow \left[\sum_{t=1}^T \frac{[Q_t(P_t - V_t) - F - A](1 - Tax) + A}{(1 + r)^t} \right] - I_0$ </div>
	<div style="margin-left: 20px;"><u>NPV</u></div>

Вертикальная линия, означающая тело программы, берётся из палитры программирования (оператор *Add line*). Внутри программы локальное присваивание

обозначается стрелкой \leftarrow из той же палитры, последний *placeholder* программы должен всегда содержать величину, возвращаемую программой.

Анализ результатов.

$$OMNPV := \text{mean}(NPV)$$

$$OMNPV = 1.237 \times 10^3$$

– оценка математического ожидания NPV ;

$$Osig := \text{Stdev}(NPV)$$

$$Osig = 970.157 \quad \text{– оценка стандартного отклонения } NPV.$$

Оценка коэффициента вариации

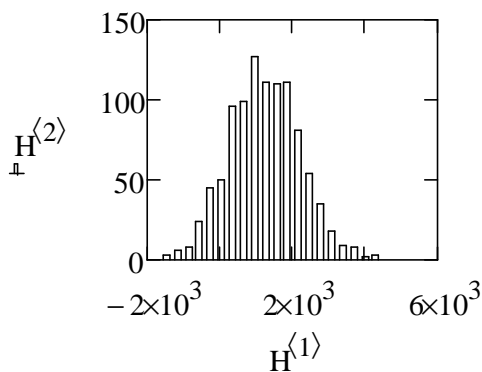
$$CV := \frac{Osig}{OMNPV}$$

.

$CV = 0.784$, подобное значение рассматривается как приемлемый риск.

$CV > 1$ считается риском выше среднего.

$H := \text{histogram}\left(\frac{N}{50}, NPV\right)$ – построение гистограммы.



По виду гистограммы можно предположить нормальное распределение NPV .

Риск инвестиционного проекта (вероятность) тогда можно рассчитать так:

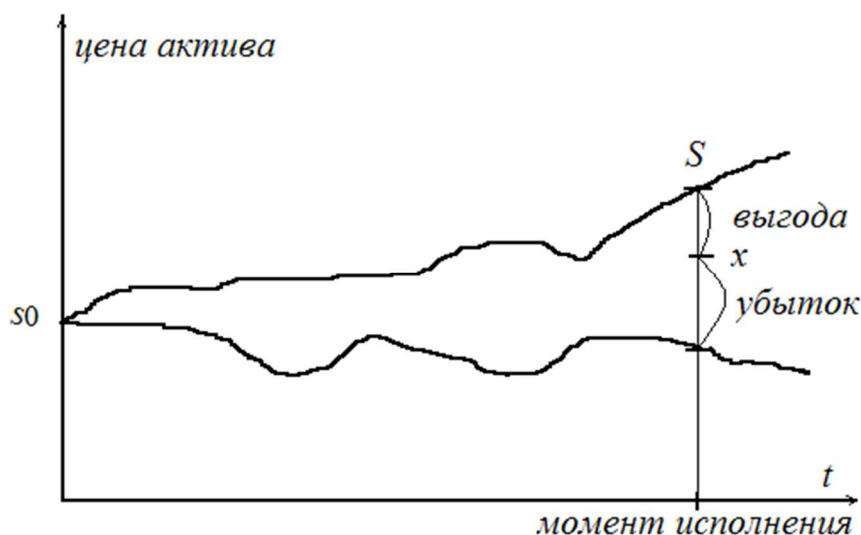
$$R := \text{pnorm}(0, OMNPV, Osig) , \quad R = 0.101$$

Если оценить вероятность частотой, то получится близкий результат

$R1 = 0.1$ – вероятность убыточности проекта невелика.

4.5. Определение цены опциона методом имитационного моделирования

Опцион на покупку – документ, дающий право, но не обязательство на покупку актива по указанной в нём цене (цена исполнения) в указанную в нём дату (момент исполнения).



Если пренебречь инфляцией, т. е. не рассматривать дисконтирование, то ясно: выгода от обладания опционом на момент исполнения (ценность опциона) – это цена актива на момент исполнения S минус цена исполнения x . Поскольку владелец опциона может отказаться от него без последующих обязательств, ценность опциона не может быть отрицательной (убытком):

$$V = \max(S - x, 0).$$

Будущая цена актива S случайна, значит, и ценность V случайна, и ценой опциона, т. е. тем, что за опцион стоит отдать в момент покупки, надо считать по определению MV – математическое ожидание ценности.

Как образуется S ? Будем считать, что цена актива непрерывно наращивается с однодневной ставкой. Пусть в первый день она была R_1 . Значит, через один день цена актива

$$S_1 = s_0 \cdot e^{R_1}. \quad (17)$$

Для одногодичного опциона (250 торговых дней) $S = S_{250} = s_0 \cdot e^{R_1 + R_2 + \dots + R_{250}}$. Для простоты считаем R_i независимыми, одинаково распределёнными величинами.

Для определения MV с заданной точностью применим имитационное моделирование: для получения конечной цены актива в n -м прогоне $s250_n = s0 \cdot e^{r_1^n + r_2^n + \dots + r_{250}^n}$ каждый раз генерируется 250 значений однодневных ставок r_t^n . При этом обычно используется эмпирическое распределение, построенное по многодневным наблюдениям за однодневной ставкой. Затем получаем N «наблюдённых путём проигрывания» модели значений ценностей опциона $v_n = \max(s250_n - x, 0)$, и оцениваем м.о. ценности как $\widehat{MV} = \bar{v}$.

4.6. Моделирование кредитных рисков

Банк назначает каждому клиенту (заёмщику):

- 1) вероятность дефолта (в планируемом горизонте – обычно один год) DP (*default probability*);
- 2) доля от займа, теряемая при дефолте LGD (*loss given default*);
- 3) сумма займа, подверженная дефолту EAD (*exposure at default*).

Случайные потери от клиента тогда $\tilde{L} = EAD \times LGD \times L$

$L = 1_D$ – бернуллиевская случайная величина (исход займа), равна 1, если событие D – дефолт свершилось, и 0 в противоположном случае. $P(D) = DP$.

Будем считать для простоты EAD и LGD неслучайными постоянными величинами. $EAD \times LGD = e$. Тогда ожидаемые потери от клиента $EL = M\tilde{L} = eM1_D = eDP$.

Калибровка DP может производиться по-разному, например, из кредитных рейтингов, выставляемых фирмам-заёмщикам специализированными агентствами: *Moody's Investors Services*, *Standard & Poor's (S&P)*, *Fitch*.

Пример 4.4. Из *Moody's* сайта или с других ресурсов легко получить доступ к их последним годовым данным о частотах корпоративных дефолтов облигаций фирм с различными кредитными рейтингами. Например, за годы 1983 по 2000 год.

<i>Rating</i>	1983	1984	...	1999	2000
<i>Aaa</i>	0.00%	0.00%	...	0.00%	0.00%
<i>Aa1</i>	0.00%	0.00%	...	0.00%	0.00%
<i>Aa2</i>	0.00%	0.00%	...	0.00%	0.00%
<i>Aa3</i>	0.00%	0.00%	...	0.00%	0.00%
<i>A1</i>	0.00%	0.00%	...	0.00%	0.00%
<i>A2</i>	0.00%	0.00%	...	0.00%	0.00%
<i>A3</i>	0.00%	0.00%	...	0.00%	0.00%
<i>Baa1</i>	0.00%	0.00%	...	0.00%	0.29%
<i>Baa2</i>	0.00%	0.00%	...	0.00%	0.00%
<i>Baa3</i>	0.00%	1.06%	...	0.34%	0.98%
<i>Ba1</i>	0.00%	1.16%	...	0.47%	0.91%
<i>Ba2</i>	0.00%	1.61%	...	0.00%	0.66%
<i>Ba3</i>	2.61%	0.00%	...	2.27%	1.51%
<i>B1</i>	0.00%	5.84%	...	3.08%	3.25%
<i>B2</i>	10.00%	18.75%	...	6.68%	3.89%
<i>B3</i>	17.91%	2.90%	...	9.90%	9.92%

Если взять средние частоты p^n по годам для каждой группы (x^n – номер группы):

	x^n	p^n	\hat{p}^n %
<i>Aaa</i>	1	<i>not</i>	0,00427
<i>Aa1</i>	2	<i>not</i>	0,00709
<i>Aa2</i>	3	<i>not</i>	0,0117
<i>Aa3</i>	4	0.08%	0,0195
<i>A1</i>	5	<i>not</i>	0,0323
<i>A2</i>	6	<i>not</i>	0,0537

A3	7	not	0,0891
Baa1	8	0.06%	0,147
Baa2	9	0.06%	0,245
Baa3	10	0.46%	0,406
Ba1	11	0.69%	0,674
Ba2	12	0.63%	1,11
Ba3	13	2.39%	1,85
B1	14	3.79%	3,07
B2	15	7.96%	5,10
B3	16	12.89%	8,46

Далее строим диаграмму рассеивания p^n от x^n , приходим к мысли об экспоненциальной модели, линеаризуем её логарифмированием, оцениваем параметры регрессии:

Regression Summary for Dependent Variable: LN-V2 (moody.s.ta) R= ,91803429 R?= ,84278695 Adjusted R?= ,82313532 F(1,8)=42,886 p = 0,000179						
	Beta	Std.Err. - of Beta	B	Std.Err. - of B	t(8)	p-level
Intercept			-5,95627	0,904451	-6,58551	0,000172
Var1	0,918034	0,140184	0,50567	0,077216	6,54877	0,000179

$$\ln \hat{p} = -5.96 + 0.506x,$$

$$\hat{p} = 2.6 * 10^{-3} e^{0.506x}.$$

Это выражение даёт последний столбец (\hat{p}^n %) для назначенных вероятностей дефолта для каждого кредитного рейтинга.

EL – это страховка или резерв для покрытия потерь банка, ожидаемых по опыту. Но такой "подушки" может оказаться недостаточно. Может быть случайное превышение против EL , *Unexpected Loss* – случайные потери. Естественная мера $UL = \sqrt{D\tilde{L}}$.

Если рассмотреть портфель из m займов $i = 1, \dots, m$: $\tilde{L}_i = e_i L_i$, $L_i = 1_{D_i}$,

$P(D_i) = DP_i = p_i$. Портфельные потери $\tilde{L}_{PF} = \sum_{i=1}^m \tilde{L}_i$.

Ожидаемые потери портфеля $EL_{PF} = M\tilde{L}_{PF} = \sum_{i=1}^m e_i p_i$, случайные потери

$$UL_{PF}^2 = D\tilde{L}_{PF} = \sum_{i=1}^m \sum_{j=1}^m e_i e_j \sqrt{p_i(1-p_i)p_j(1-p_j)} \rho_{ij} \quad (17.1)$$

где $\rho_{ij} = \rho(L_i, L_j)$ – коэффициент корреляции исходов займа заёмщиков i и j .

.....

Пример 4.5. Портфель из двух займов.

$$UL_{PF}^2 = e_1^2 p_1(1-p_1) + e_2^2 p_2(1-p_2) + 2e_1 e_2 \sqrt{p_1(1-p_1)p_2(1-p_2)} \rho.$$

1. $\rho = 0$ (совершенная диверсификация). Исходы займов независимы. UL_{PF} малы.

2. $\rho > 0$, т. е. дефолт одного повышает вероятность дефолта другого. Действительно,

$$\begin{aligned} P\{L_2 = 1 | L_1 = 1\} &= \frac{P\{L_2 = 1, L_1 = 1\}}{P\{L_1 = 1\}} = \frac{ML_2 L_1}{p_1} \\ &= \frac{p_1 p_2 + \text{cov}(L_2, L_1)}{p_1} = p_2 + \frac{\text{cov}(L_2, L_1)}{p_1}. \end{aligned} \quad (17.1.1)$$

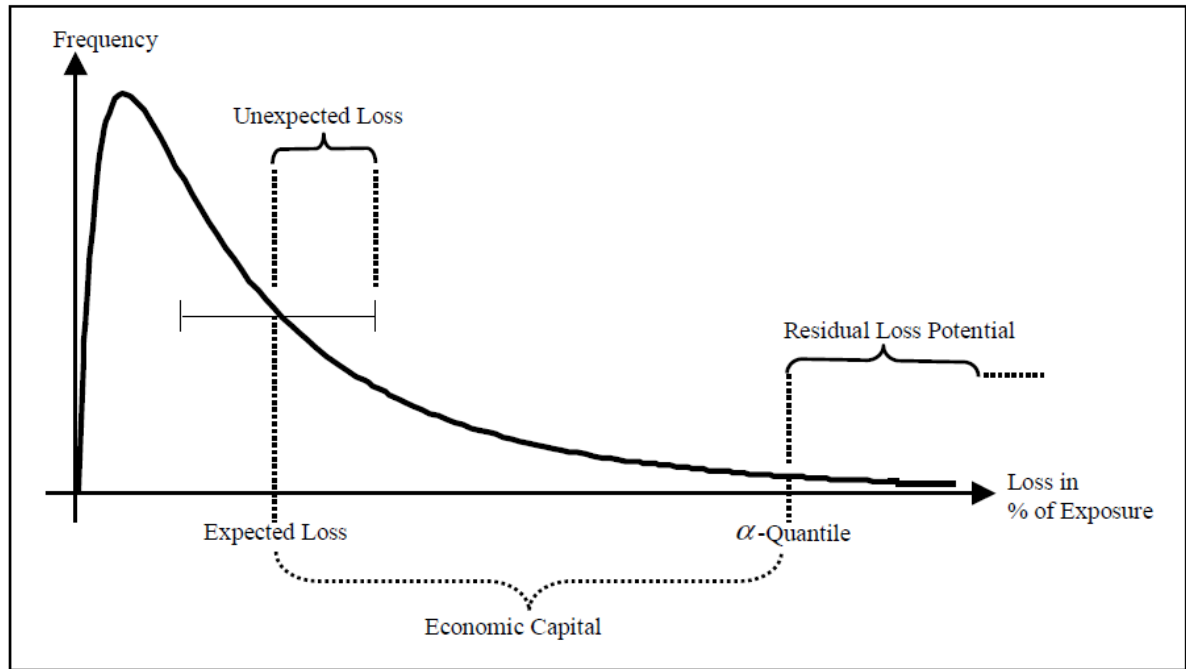
....

Это $> p_2$ при $\rho > 0$. При $\rho = 1$ и $p_1 = p_2 = p$, $e_1 = e_2 = e$, $UL_{PF} = 2e\sqrt{p(1-p)}$ т. е. удвоенный риск одного заёмщика (concentration risk).

3. $\rho < 0$ и если возможно достичь -1, то заём 2 есть perfect hedge против займа 1 и $UL_{PF} = 0$.

Экономичный капитал (EC – economic capital)

И дополнительной подушки UL_{PF} может оказаться недостаточно (бывают отклонения и в 2 и 3 ... сигма). Уточнить величину этого капитала риска можно введя понятие EC (или Capital at Risk (CaR) или кредитное Value-at-Risk (VaR)).



$$EC_{\alpha} = q_{\alpha} - EL_{PF} .$$

Пример 4.6. Если $\alpha = 0.99$, то EC_{α} будет (в среднем) достаточно в 99 годах из 100 для покрытия неожиданных потерь.

$EL_i = e_i p_i$ перекладываются на заёмщика, $(EC_{\alpha})_i$ тоже перекладывается на заёмщика как доля EC_{α} портфеля, добавляемая за счёт нового займа.

....

Распределение портфельных потерь

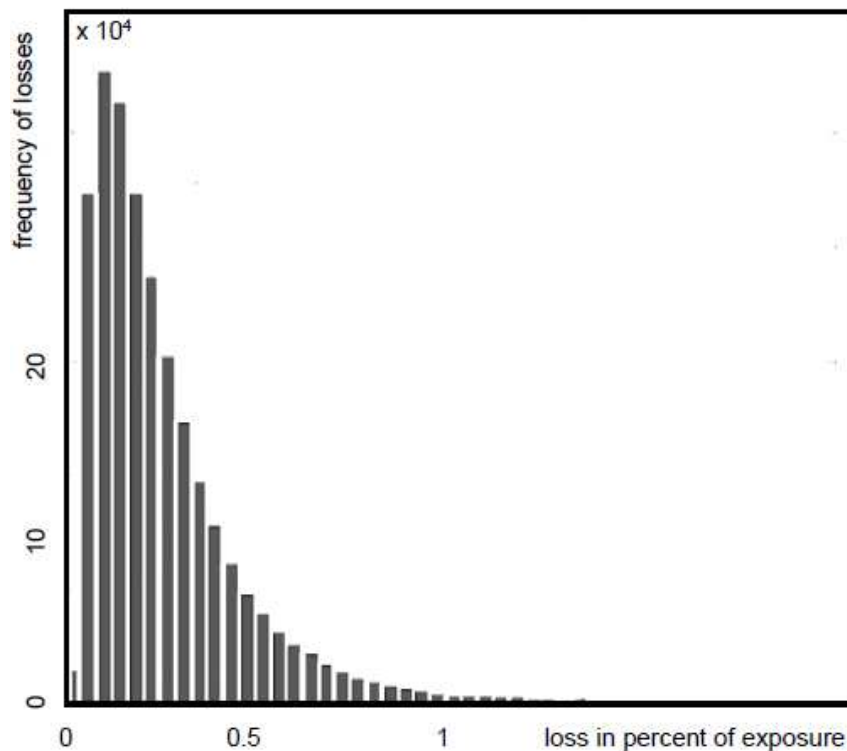
Есть два пути получить распределение портфельных потерь \tilde{L}_{PF} : имитационное моделирование и аналитическая аппроксимация.

ИМ. Пусть мы сгенерировали

$$\{\tilde{l}_{PF}^1, \tilde{l}_{PF}^2, \dots, \tilde{l}_{PF}^N\} \quad (17.2)$$

– возможные потери портфеля, учтя распределения потерь отдельных займов и их корреляции. Тогда по этой выборке можно построить эмпирическую функцию распределения портфельных потерь $\hat{F}_N(x)$ или эмпирическую плотность распределения (гистограмму).

Типичный вид гистограммы (портфель 2000 корпоративных займов среднего размера):



Для оценки экономического капитала нужна оценка квантили порядка α :

$$\hat{q}_\alpha = \begin{cases} \tilde{l}_{PF}^{(ceil(N\alpha))}, & N\alpha - \text{дробное}; \\ \alpha \tilde{l}_{PF}^{(N\alpha)} + (1 - \alpha) \tilde{l}_{PF}^{(N\alpha+1)}, & N\alpha - \text{целое}; \end{cases}$$

где $\{\tilde{l}_{PF}^{(n)}\}$ – вариационный ряд выборки (17.2). Тогда $\widehat{EC}_\alpha = \hat{q}_\alpha - \frac{1}{N} \sum_{n=1}^N \tilde{l}_{PF}^n$.

4.7. Упражнения и лабораторный практикум

Упражнение 4.1.

При 100 прогонах имитационной модели среднее выходной характеристики составило 22, оценка стандартного отклонения выходной характеристики составила 3.56. Найти приближенный доверительный интервал для математического ожидания выходной характеристики с доверительной вероятностью 0.997. Ответ 20.9 23.1.

Упражнение 4.1a.

После 100 прогонов имитационной модели достигнута точность определения результата 0.02. Сколько надо провести прогонов для достижения точности 0.005? Ответ 1600.

Упражнение 4.2.

Повторить в MathCAD пример 4.1. Найти математическое ожидание Y с точностью 0.5% (с вероятностью 95%).

Упражнение 4.3.

$$I = \int_0^{\pi/2} \sqrt{\sin(x)} dx$$

Вычислить интеграл I методом Монте-Карло с точностью 0.5% (с вероятностью 95%). Проверить, что доверительный интервал с должной частотой покрывает истинное значение 1.19814... (рассмотреть случаи 2, 1 и (?) сигма). К чему приведёт изменение N ?

Упражнение 4.3a

Методом Монте-Карло, используя случайные числа $\sim U(0, 1)$ 0.1, 0.09, 0.73, 0.45, 0.33 найти приближённое значение интеграла $\int_0^{1.2} \sqrt{\cos(x)} dx$. Ответ 1.12.

Найти стандартную ошибку результата. Ответ 0.0437.

Упражнение 4.4.

Вычислить интеграл методом Монте-Карло с точностью 0.5% (с вероятностью 95%).

$$\iiint_{[0,1]^3} \frac{x^2 y dx dy dz}{\left[1 + x^2 \left[1 + y^2 (1 + z^2)\right]\right]^2}$$

Заменить поэлементное вычисление выборки значений интеграла по теореме о среднем – векторным, с помощью оператора Vectorize. Попробовать убрать стрелку. Прodelать это же в предыдущем упражнении. Сделать выводы.

Упражнение 4.5.

Вычислить интеграл методом Монте-Карло с точностью 0.5% (с вероятностью 95%).

$$\iiint_{[0,2]^3} \frac{\ln(1+x^2)}{\exp(yz)(3+\cos(x+z))} dx dy dz$$

Упражнение 4.6.

Повторить пример 4.3. Найти математическое ожидание NPV с точностью 0.5% (с вероятностью 95%). Оценить риск долей наблюждённых NPV, меньших нуля. Оценить коэффициент корреляции NPV и средних за время проекта переменных затрат VAV. Оценить линейную функцию регрессии NPV на VAV и построить её график на фоне корреляционного поля (диаграммы рассеивания). Дать прогноз NPV при средних за время проекта переменных затратах в 37 у. е.

Упражнение 4.6a

Пусть NPV имеет нормальное распределение со стандартным отклонением 100 у.е. Коэффициент вариации = 1.2. Каков риск инвестиционного проекта (вероятность получить убыток)? Ответ 0.202.

Упражнение 4.6б

Пусть NPV имеет нормальное распределение с математическим ожиданием 10 у.е. и дисперсией 56.25. Каков риск инвестиционного проекта (вероятность получить убыток)? Ответ 0.0912.

Упражнение 4.7.

При начальной цене актива $s_0 = 1000$ и цене исполнения $x = 1100$ найти цену одногодичного опциона (250 торговых дней) с точностью 1% (с вероятностью 0.95). Для однодневной ставки непрерывного наращения R принять $\sim N(0.000455, 0.01)$. Как изменится цена опциона, если цену исполнения x установить равной 1000?

Упражнение 4.7а.

Вывести формулу (17).

Упражнение 4.8.

Вывести формулы (17.1), (17.1.1).

Упражнение 4.9.

Получить выражение для $P\{L_2 = 0 \mid L_1 = 0\}$.

Упражнение 4.9а

Априорные вероятности дефолтов двух заёмщиков $p_1 = 0.3$ и $p_2 = 0.4$ коэффициент корреляции $\rho = 0.6$. Найти вероятность выжить второму, при условии дефолта первого. Ответ 0.151.

Упражнение 4.9б

Априорные вероятности дефолтов двух заёмщиков $p_1 = 0.3$ и $p_2 = 0.4$ коэффициент корреляции $\rho = 0.6$. Найти вероятность выжить второму, при условии выживания первого. Ответ 0.792.

Упражнение 4.9г

Априорные вероятности дефолтов двух заёмщиков $p_1 = 0.3$ и $p_2 = 0.4$ коэффициент корреляции $\rho = 0.6$. Найти вероятность дефолта и второго и первого. Ответ 0.255.

Упражнение 4.9д

Априорные вероятности дефолтов двух заёмщиков $p_1 = 0.3$ и $p_2 = 0.4$ коэффициент корреляции $\rho = 0.6$. Найти вероятность дефолта второго и выживания первого. Ответ 0.145.

Упражнение 4.10.

Для кредитного портфеля из двух займов размерами $e_1 = 100$ у.е. (рейтинг *Moody's* заёмщика B2 5,10%) и $e_2 = 200$ (B3 8,46%)

1) сгенерировать $N = 1000$ скоррелированных (с $\rho = 0.7$) пар исходов займов.

Для этого:

- а) сгенерировать N исходов l_1 1-о заёмщика,
- б) $n = 1$,
- в) если $l_{1n} = 1$, генерировать l_{2n} из распределения $p(l_2 | l_1 = 1)$, иначе генерировать l_{2n} из распределения $p(l_2 | l_1 = 0)$,
- г) $n = n + 1$,
- д) если $n > N$ то *Stop*, иначе переход к пункту в);

2) для проверки оценить коэффициент корреляции и МО, построить диаграмму рассеивания;

3) получить $N = 1000$ потерь портфеля (в процентах), построить гистограмму, оценить ожидаемые потери (верифицировать) и случайные потери (верифицировать);

4) оценить 97% экономичный капитал;

5) сравнить всё со случаями $\rho = 0$, -0.07 .

Переделать программу на случай портфеля m независимых займов.

Упражнение 4.11.

Априорные вероятности дефолтов двух заёмщиков $p_1 = 0.3$ и $p_2 = 0.4$, коэффициент корреляции $\rho = 0.6$. Используя последовательно псевдослучайные числа $\sim U(0, 1)$: 0.34, 0.78, 0.22, 0.55, 0.11, 0.85 сгенерировать 3 пары исходов займов. Для этого: а) сгенерировать 3 исхода 1-о заёмщика, б) для каждой из

них с помощью условного распределения сгенерировать исход 2-о заёмщика. Всегда при генерации вероятность дефолта откладывать первой (от 0). В ответе привести последовательно 3 полученные пары. Ответ 0 0 0 1 1 0.

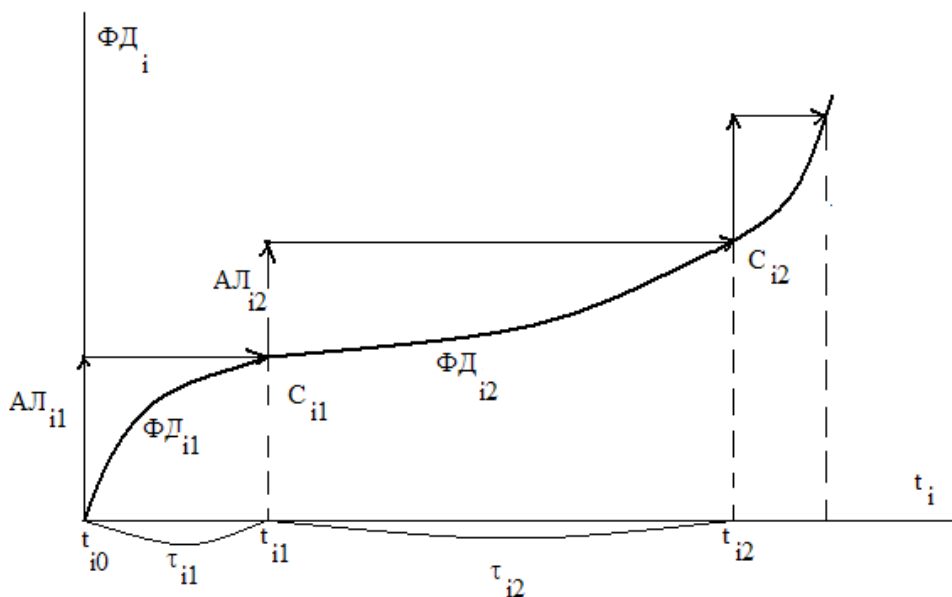
Упражнение 4.12.

Применить в упр. 4.10 для генерации бернуллиевских коррелированных величин метод упр. 3.7.

5. Общие принципы имитационного моделирования многокомпонентных систем

Система – совокупность элементов, объединённых некоторой формой регулярного взаимодействия или взаимозависимости для выполнения заданной функции. Отдельные элементы системы или её подсистемы называются компонентами.

Пример 5.1. Движение эскадрильи самолётов на учениях. Каждый i -й самолёт – компонента системы, в его движении можно выделить ряд стадий, в ходе которых выполняется последовательность функциональных действий (ФД). Для каждой i -й компоненты вводится своё локальное время t_i . При t_{i0} начинается разбег и взлёт i -го самолёта со взлётной полосы – это функциональное действие $\Phi Д_{i1}$. Боевой разворот – $\Phi Д_{i2}$. Стрельба по мишени – $\Phi Д_{i3}$ и т. д.



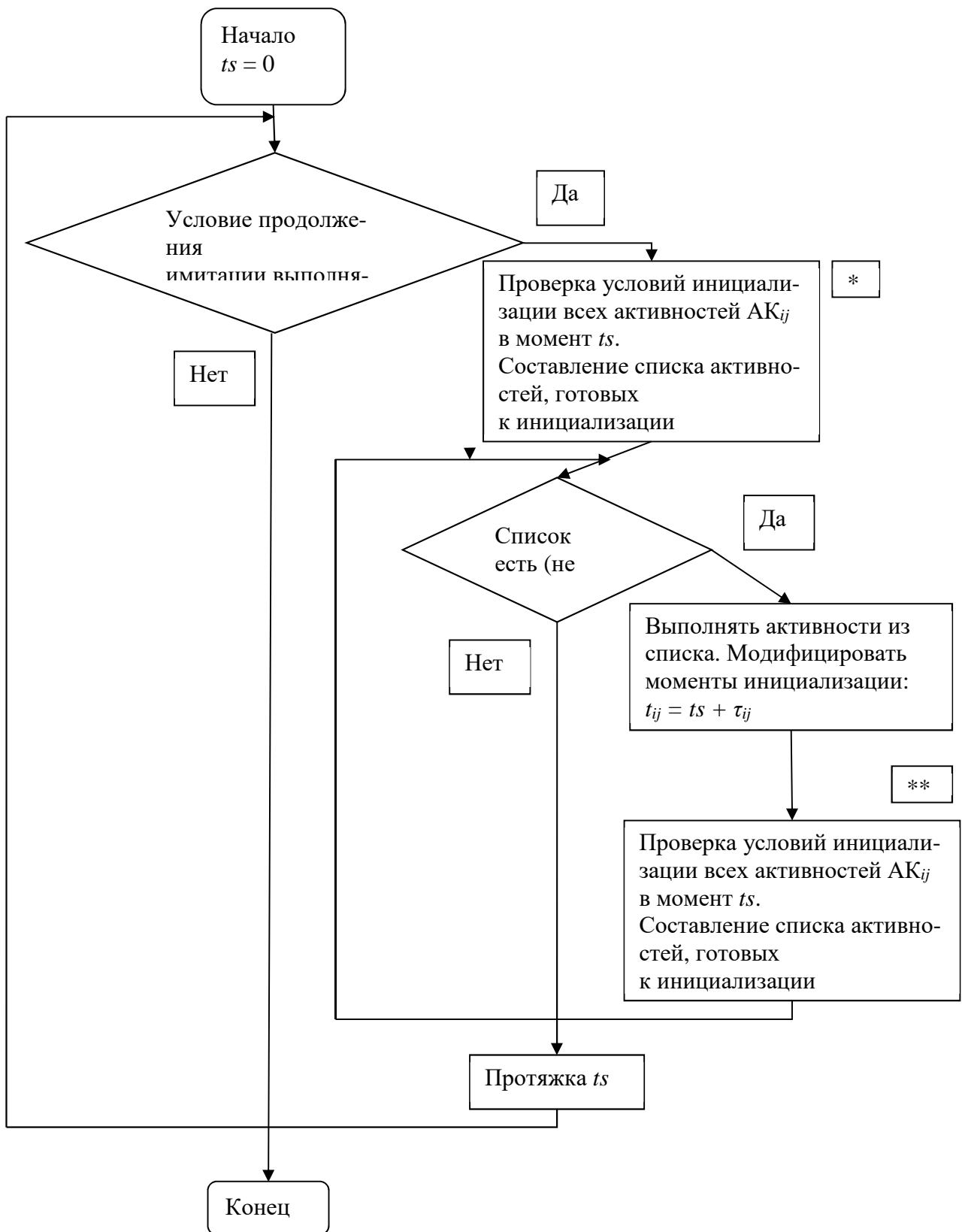
Любое $\Phi Д_{ij}$ выполняется на интервале τ_{ij} и завершается событием C_{ij} .

В имитационной модели $\Phi Д_{ij}$ аппроксимируется некоторым упрощённым функциональным действием (алгоритм $АЛ_{ij}$), который реализуется при неизменном значении t_i , а затем осуществляется изменение t_i на τ_{ij} , инициируя появление события C_{ij} . Пара $(АЛ_{ij}, \tau_{ij}) = АК_{ij}$ – активность (работа).

Если бы на компьютере имитировалось поведение только одной компоненты, то выполнение активностей можно было бы осуществлять строго последовательно, пересчитывая каждый раз t_i . На самом деле компонент много и функционируют они одновременно, а в большинстве компьютеров в каждый конкретный момент может исполняться алгоритм только одной из компонент модели. Необходимо реализовать квазипараллелизм – выстраивание параллельно происходящих активностей в последовательный порядок, с учётом их должной синхронизации (для учёта взаимовлияния). Это делается с помощью глобальной переменной ts – системного времени. Системное время – это мыслимое (в часах, днях, годах) модельное время, в котором происходит эволюция модели. Фактически это последовательность моментов, в которых нужно выполнять какие-либо активности компонент.

5.1. Организация квазипараллелизма просмотром активностей

Блок-схема моделирования:



Примечания к схеме:

* – в простейшем случае это проверка по совпадению времени инициализации с ts ;

** – новая проверка необходима, т. к. *начало* выполнение некоторых активностей может открыть условия инициализации других активностей в тот же момент ts . Например, в момент ts 1-й самолёт готов начать боевой разворот AK_{12} . 2-й самолёт в тот же момент в принципе готов к разбегу и взлёту (AK_{21}), но пока предыдущий не начал боевого разворота (не ушёл из зоны взлётной полосы) – нельзя! Поэтому в список * AK_{21} не попадёт (там будет AK_{12}), а в список ** попадёт. В итоге в ts оба начнут свои действия одновременно, но процессор выполнит их последовательно.

5.2. Два способа изменения (протяжки) системного времени

1. До следующего ближайшего момента инициализации (синхронно-событийный способ):

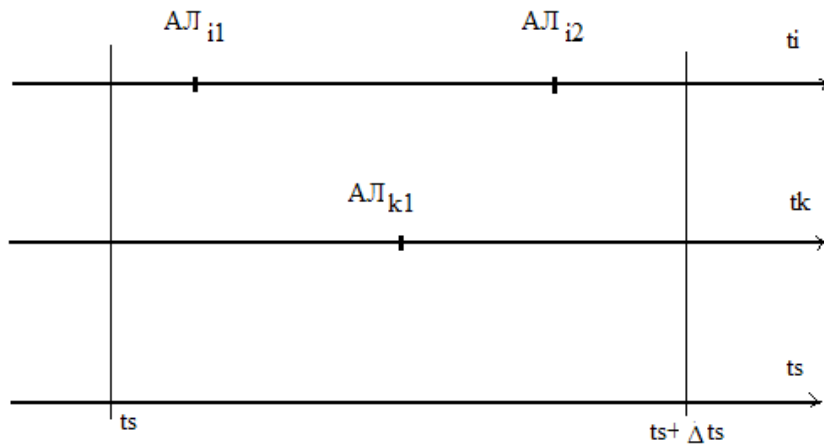
$$ts = \min(t_{ij}).$$

2. Способ фиксированного шага (синхронно-временной):

$$ts = ts + \Delta ts, \quad \Delta ts = \text{const},$$

и все активности с временами инициализации, попадающими между старым моментом ts и новым $ts + \Delta ts$, считаются готовыми к инициализации и как бы одновременными. Их выполнение происходит в неопределённом порядке. Ясно, что при этом могут возникать ошибки моделирования.

Пример 5.2. СМО с одним местом в очереди. $АЛ_{i1}$, $АЛ_{i2}$ – алгоритм источника: появление заявки, постановка её в очередь или потеря, если места в очереди нет. $АЛ_{k1}$ – алгоритм канала обслуживания: готовность принять заявку.



На самом деле потери клиента быть не должно, но если в момент $ts + \Delta ts$ алгоритмы выполнять так: AL_{i1} , AL_{i2} , AL_{k1} , то AL_{i2} уже вызовет потерю. Поэтому желательно Δts брать поменьше. Синхронно-временной способ применяется, когда события появляются через фиксированные промежутки времени или когда их много и они появляются группами. Во всех остальных случаях применяется синхронно-событийный способ.

Пример 5.3. СМО $(M|M|1|\infty)$. В таком четырёхпольном обозначении первое поле характеризует распределение интервалов времени между заявками источника (M – экспоненциальное, т. е. это пуассоновский поток); второе поле характеризует распределение времён выполнения заявок каналами обслуживания; третье – число каналов; четвёртое – число мест в очереди.

Пусть интенсивность потока заявок $la = 9$ заявок/час; интенсивность обслуживаний $mu = 12$ заявок/час; время моделирования $tsmax = 100$ час.

```

k := | i ← 0
      | ts ← 0
      | tk ← ts
      | ti ← ts
      | q ← 0
      | s ← 0
      | z ← 0
      | while ts < tsmax
        |   in ← g(ts, ti, tk)
        |   while (in1 = 1) + (in2 = 1)
        |     | if in1 = 1
        |       |   q ← q + 1 if i > 0
        |       |   i ← i + 1
        |       |   ti ← ti +  $\frac{-1}{\text{la}} \cdot \ln(1 - \text{rnd}(1))$ 
        |       | if in2 = 1
        |       |   | if q > 0
        |       |     | s ← s + 1 if z = 1
        |       |     | q ← q - 1
        |       |     | z ← 1
        |       |     | tk ← tk +  $\frac{-1}{\text{mu}} \cdot \ln(1 - \text{rnd}(1))$ 
        |       |   otherwise
        |       |     | if z = 1
        |       |       | s ← s + 1
        |       |       | z ← 0
        |       |     | tk ← tsmax
        |       |   in ← g(ts, ti, tk)
        |       | ts ← min[(ti tk)T]
        |       | tk ← ts if tk = tsmax
        |       |  $k^{(i)} \leftarrow \begin{pmatrix} \text{ti} \\ \text{tk} \\ q \\ s \end{pmatrix}$ 
        |   kT

```

$$g(ts, ti, tk) := \begin{cases} ii \leftarrow 1 & \text{if } ti = ts \\ ii \leftarrow 0 & \text{otherwise} \\ ik \leftarrow 1 & \text{if } tk = ts \\ ik \leftarrow 0 & \text{otherwise} \\ \begin{pmatrix} ii \\ ik \end{pmatrix} \end{cases}$$

Это подпрограмма проверки условий инициализации активностей и создания списка.

Активность источника. Если это не нулевая заявка (которой фактически не было), то ставит её в очередь, учитывает число заявок, модифицирует своё локальное время появления следующей заявки.

Активность канала. Если очередь не пуста, если канал был занят, то выпуск обслуженной заявки, впуск новой, канал – в состояние «занят», модифицирует своё локальное время окончания обслуживания. Если очередь была пуста, если канал был занят, то выпуск обслуженной, канал – «свободен», время очередной активизации – по активизации источника. Пока его устанавливаем равным ts_{max} , чтобы не ставить понапрасну в список инициализации в то же время.

$k =$

	1	2	3	4
1	0.00014	0.00014	0	0
2	0.02401	0.07343	0	0
3	0.07193	0.07343	1	0
4	0.26423	0.26423	0	3
5	0.30449	0.30449	0	4
6	0.3222	0.67667	0	4
7	0.33629	0.67667	1	4

	1	2	3	4
531	54.60179	54.64143	16	513
532	54.96226	54.99505	13	517
533	54.96817	54.99505	14	517
534	55.48982	55.5337	11	521
535	55.64733	55.64745	11	522
536	55.69742	55.7914	11	523
537	55.76079	55.7914	12	523

k =

k =

	1	2	3	4
919	99.20751	99.35135	0	917
920	99.30711	99.35135	1	917
921	99.40969	99.41102	0	919
922	99.6129	99.6129	0	921
923	99.83404	99.83404	0	922
924	99.83778	99.90466	0	922
925	100.10673	100	0	...

$x := k^{(3)}$

$oMq := \text{mean}(x)$

$osig := \text{Stdev}(x)$

$oMq = 2.98703$

$osig = 3.44654$

В результате имитации получена оценка средней длины очереди 2.987 заявок. Этот результат будет обсуждаться в разделе 6.4. Оценка стандартного отклонения длины очереди – 3.447 заявок.

Эту программу легко обобщить на моделирование системы $(reg|M|1|\infty)$, с регулярным потоком заявок и $(M|M|1|m)$, с конечным числом мест в очереди, а также на протяжку системного времени синхронно-временным способом.

5.3. Организация квазипараллелизма транзактным способом

Этот способ особенно удобен для имитации систем массового обслуживания (СМО). Каждая поступившая заявка (транзакт) образует компоненту системы, для неё устанавливается локальная временная координата t_i . Активность проявляют здесь не приборы обслуживания (каналы), а сами транзакты, занимая тот

или иной свободный прибор и модифицируя своё время следующей инициализации: $t_{ij} = ts + \tau_{ij}$.

Существуют специальные языки моделирования транзактным способом, например *GPSS (General Purpose System Simulator)*.

Пример 5.4. Моделирование в *GPSS* той же системы, что и в примере 5.3.

Barber STORAGE 1 – описание *Barber* как 1 параллельно работающий канал.

GENERATE (Exponential(1,0,1/9)) – генерация транзакта через случайный промежуток времени с экспоненциальным распределением, 1 – *seed*, 0 – сдвиг, 1/9 – среднее время между заявками.

QUEUE Hall запись транзакта в очередь с именем *Hall*

ENTER Barber вход транзакта в канал

DEPART Hall уход из очереди

ADVANCE (Exponential(1,0,1/12)) задержка в канале (обслуживание)

LEAVE Barber освобождение канала

TERMINATE удаление транзакта

GENERATE 100 моделирование 100 единиц времени

TERMINATE 1 конец моделирования

Написав программу, идём в меню *Command*→*Create Simulation*, *Command*→*Start [Start1]*→*OK*.

Отчёт:

<i>START TIME</i>	<i>END TIME</i>	<i>BLOCKS</i>	<i>FACILITIES</i>	<i>STORAGES</i>
0.000	100.000	9	0	1

<i>NAME</i>	<i>VALUE</i>
<i>BARBER</i>	10000.000
<i>HALL</i>	10001.000

<i>LABEL</i>	<i>LOC</i>	<i>BLOCK TYPE</i>	<i>ENTRY COUNT</i>	<i>CURRENT COUNT</i>	<i>RETRY</i>
	1	<i>GENERATE</i>	922	0	0
	2	<i>QUEUE</i>	922	0	0
	3	<i>ENTER</i>	922	0	0
	4	<i>DEPART</i>	922	0	0
	5	<i>ADVANCE</i>	922	1	0
	6	<i>LEAVE</i>	921	0	0
	7	<i>TERMINATE</i>	921	0	0
	8	<i>GENERATE</i>	1	0	0
	9	<i>TERMINATE</i>	1	0	0

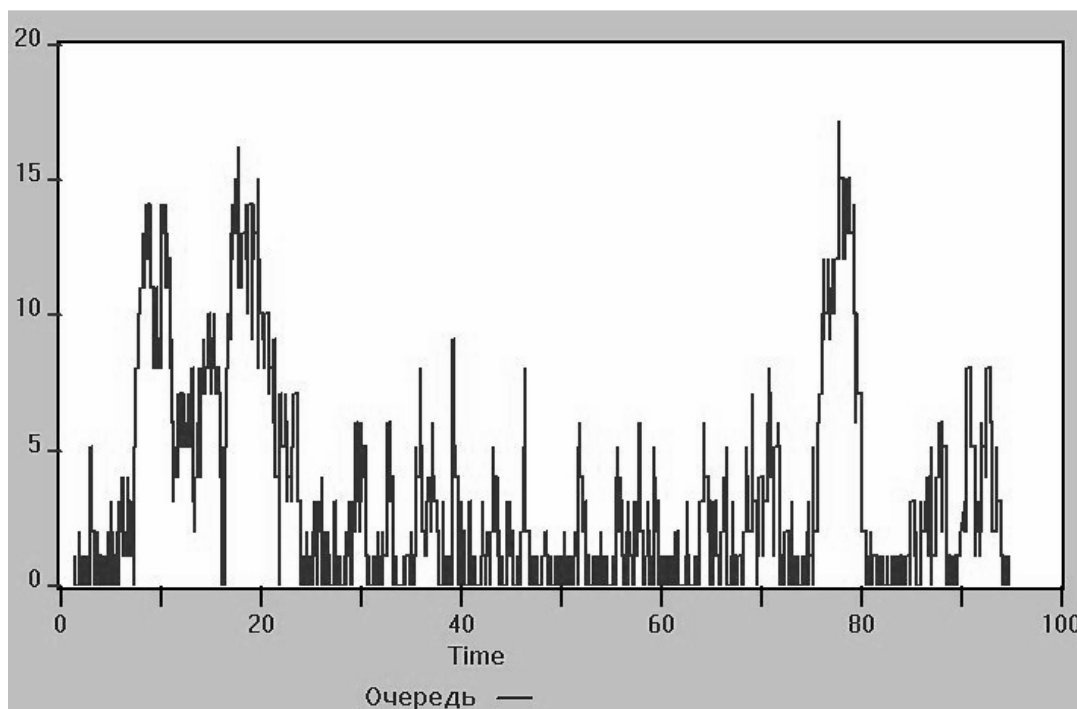
QUEUE	MAX	CONT.	ENTRY	ENTRY(0)	AVE.CONT.	AVE.TIME	AVE.(-0)	RETRY
HALL	17	0	922	192	3.133	0.340	0.429	0

STORAGE	CAP.	REM.	MIN.	MAX.	ENTRIES	AVL.	AVE.C.	UTIL.	RETRY	DELAY
BARBER	1	0	0	1	922	1	0.813	0.813	0	0

В очередь вошло 922 заявки, максимальное число заявок в очереди 17. Связь между средним числом заявок в очереди 3.133 и средним временем ожидания 0.340, будет обсуждаться в разделе 6.4. Среднее число заявок в канале обслуживания 0.813.

Для получения следующего прогона *Command* → *Reset* → *Start*, если нужно *seed* сбросить в начальное состояние *Command* → *Retranslate*.

Можно добавить вывод графика очереди *Window* → *Simulation Window* → *Plot Window*. *Label* Очередь, *Expression* *Q\$Hall* → *Plot*, *Command* → *Start*. Аналогично – график занятости канала.



5.4. Упражнения и лабораторный практикум

Упражнение 5.1.

Повторить пример 5.3 по моделированию СМО ($M|M|1|\infty$). Добавить к результатам моделирования график временной зависимости длины очереди и гистограмму. Проверить условие существования стационарного режима. Переделать программу на регулярный входной поток с той же интенсивностью заявок. Сравнить результаты со случаем пуассоновского входного потока. Рассмотреть случай регулярного обслуживания с интенсивностью 12, затем 8. Какая должна быть очередь (провести верификацию)?

Упражнение 5.2.

Нарисовать детальную блок-схему программы из примера 5.3. Обобщить её на случай конечного числа мест в очереди.

Упражнение 5.3.

Переделать пример 5.3 на синхронно-временной способ протяжки системного времени. Убедиться в ошибках моделирования. В пределе малых Δt_s получить результаты примера 5.3.

Упражнение 5.4.

Обобщить пример 5.3 на случай конечного числа мест в очереди. Сравнить результаты моделирования с дедуктивным решением (в частности, при $m = 5$ получить долю потерянных заявок вблизи 0.051). Обобщить программу на $m = 0$.

Упражнение 5.5.

Обобщить пример 5.3 на случай двух каналов. Сравнить результаты моделирования с дедуктивным решением.

6. Испытание и эксплуатация имитационных моделей

После того как имитационная модель реализована на компьютере и работает, необходимо выполнить следующие этапы.

6.1. Верификация и проверка адекватности модели

Верификация

– это проверка соответствия алгоритма функционирования имитационной модели тому замыслу, который был заложен при её разработке. Чаще всего это делается рассмотрением граничных (предельных) случаев (значений параметров), при которых имеется очевидное или дедуктивно найденное решение. Стохастические элементы заменяются на детерминированные и/или вносят другие упрощения.

Проверка адекватности модели

– сравнение имитационной модели с реальной системой, которую она моделирует, т. е. система должна существовать, что бывает не всегда.

Один из способов проверки – сравнение математических ожиданий откликов реальной системы Y^* и имитационной модели Y при одинаковых внешних воздействиях, т. е. проверяется гипотеза:

$$H_0 : MY^* = MY$$

$$H_1 : MY^* \neq MY .$$

Проводится небольшое число N_1 дорогостоящих экспериментов над реальной системой. Получаем выборку: $y_1^*, \dots, y_{N_1}^*$. Проводится большее число N_2 экспериментов над имитационной моделью. Получаем выборку: y_1, \dots, y_{N_2} . Вычисляется

уровень значимости данных

$$sl = 2 \left(1 - \Phi_{N_1+N_2-2}^{Cm} \left(\left| \frac{\bar{y} - \bar{y}^*}{s_{\Sigma} / \sqrt{\frac{N_1 N_2}{N_1 + N_2}}} \right| \right) \right), \quad \text{где}$$

$$s_{\Sigma}^2 = \frac{(N_1 - 1)s_*^2 + (N_2 - 1)s^2}{N_1 + N_2 - 2}. \text{ Подобную проверку необходимо провести для}$$

всех фиксируемых откликов.

6.2. Оценка погрешности результирующего показателя имитации из-за различия затравочных чисел генератора псевдослучайных чисел

Генераторы псевдослучайных чисел неидеальны и использование различных «затравочных», *seed* чисел приводит к различным последовательностям псевдослучайных чисел, а значит и к различным откликам.

Или, если генератор считать вполне случайным, то это разброс отклика из-за случайной природы «первотолчка» (начальных условий работы системы).

В любом случае погрешности результирующего показателя имитации из-за различия *seed* можно оценить и уменьшить, проведя *M* перезапусков имитации при различных *seed*.

Пример 7.1. Мы считаем, что есть “истинный” результат (*result*) а наблюдения отклоняются от него из-за ошибок, связанных с *seed*. Тогда $MRES = result$. Пусть *res* – вектор-выборка результирующего показателя имитационного моделирования *RES*. Получена при различных *seed* при прочих равных условиях.

$$res := (2 \ 9 \ 6 \ 8 \ 5 \ 3 \ 6 \ 6 \ 4 \ 3)^T$$

$$M := \text{length}(res)$$

$$M = 10$$

$oresult := \text{mean}(res)$ – оценка результата как оценка математического ожидания по выборке

$$osRES := \text{Stdev}(res) \text{ – оценка стандартного отклонения } RES$$

$$oresult = 5.2$$

$$osRES = 2.251$$

$$\Delta := \frac{osRES}{\sqrt{M}} qt\left(1 - \frac{1-\gamma}{2}, M-1\right) \text{ – полуширина доверительного интервала с до-}$$

верительной вероятностью γ для результата моделирования, искажённого из-за

разных *seed*. Здесь $qt(p, N)$ – квантиль порядка p распределения Стьюдента с N степенями свободы, поэтому, при $\gamma = 0.95$ (95 % доверительный интервал)

$$oresult - \Delta = 3.59 < result < oresult + \Delta = 6.81$$

Или, другими словами, $\Delta = 1.61$ погрешность и итоговый результат для *result*:

$$result = oresult \pm \Delta = 5.2 \pm 1.61.$$

Пример 7.2. Если в примере 4.1 произвести $M = 10$ перезапусков с различными *seed* (каждый раз по $N = 500$), то получим результат

$MY = (2.04 \div 2.09) \pm 2 \cdot (0.028 \div 0.0297)$ вместо $MY = 2.07 \pm 2 \cdot 0.03$ ранее. Т.е. все результаты моделирования превратились в интервалы, связанные с влиянием *seed*.

6.3. Методы понижения дисперсии

Антитетический метод

Каждый раз, когда генерируется случайное число $r \sim U(0, 1)$, вычисляется его дополнение $(1 - r)$, которое используется для параллельного определения результата. Значит, если входная переменная довольно велика, в параллельном расчёте она будет довольно мала. Как правило, это ведёт к значениям откликов Y_1 и Y_2 , которые отрицательно коррелированы. Делаем N прогонов, получаем парные результаты: $\begin{pmatrix} y_1^1 \\ y_2^1 \end{pmatrix}, \dots, \begin{pmatrix} y_1^N \\ y_2^N \end{pmatrix}$. $\{y_1^n\}$ – выборка из Y_1 , $\{y_2^n\}$ – выборка из Y_2 . Ясно, что Y_1, Y_2 распределены одинаково с Y – результат простой (без параллельных расчётов). Введём случайную величину $Y_0 = \frac{Y_1 + Y_2}{2}$.

$MY_0 = 0.5MY_1 + 0.5MY_2 = MY$, поэтому MY будем искать так:

$$MY = MY_0 = \bar{y}_0 \pm (2 \div 3) \frac{\hat{\sigma}_{Y_0}}{\sqrt{N}}. \quad (24)$$

$$DY_0 = D0.5(Y_1 + Y_2) = 0.25D(Y_1 + Y_2) = 0.25(DY_1 + DY_2 + 2\text{cov}(Y_1, Y_2)) = \\ = 0.25(DY_1 + DY_2 + 2\rho(Y_1, Y_2)\sigma_{Y_1}\sigma_{Y_2}),$$

поскольку $\rho(Y_1, Y_2)$ отрицателен, DY_0 , будет меньше DY , вплоть до $0.25(\sigma_{Y_1} - \sigma_{Y_2})^2 = 0$ при $\rho(Y_1, Y_2) = -1$. Значит, работая с Y_0 , мы имеем величину с тем же математическим ожиданием, но с уменьшенной дисперсией; поэтому оценка $\hat{\sigma}_{Y_0}$, а значит и погрешность в (24), будет мала.

Понижение дисперсии при вычислении интегралов

Пусть нужно вычислить интеграл

$$I = \int_a^b f(x)dx = \int_a^b f(x) \frac{1}{p_\xi(x)} p_\xi(x) dx = M \left(\frac{f(\xi)}{p_\xi(\xi)} \right) = M\eta;$$

здесь $p_\xi(x)$ – плотность распределения произвольной случайной величины ξ , принимающей значения на (a, b) . Далее рассуждаем, как и прежде:

$$I = M\eta = \bar{y} \pm (2 \div 3) \frac{\hat{\sigma}_\eta}{\sqrt{N}}, \text{ где } y_n = \frac{f(x_n)}{p_\xi(x_n)} - \text{выборка из } \eta, \text{ которая получится когда}$$

x_n – выборка из ξ .

Выбирая различные распределения $p_\xi(x)$ а значит и η , будем иметь различные σ_η , а значит и погрешности вычисления I . Если взять $p_\xi(x) = kf(x)$ (считаем $f(x)$

> 0), то $D\eta = D \left(\frac{f(\xi)}{p_\xi(\xi)} \right) = D \left(\frac{1}{k} \right) = \frac{1}{k^2} D1 = 0$! Коэффициент k надо находить из

условия нормировки: $1 = \int_a^b p_\xi(x) dx = k \int_a^b f(x) dx$, что дает $k = \left(\int_a^b f(x) dx \right)^{-1}$; но вхо-

дящий сюда интеграл и есть неизвестный нам искомый интеграл! Поэтому практически берут $p_\xi(x)$ сходной с $f(x)$, но так, чтобы с нормировкой всё было просто и генерация x_1, \dots, x_N из этого распределения $p_\xi(x)$ была бы несложной; статистически оценённая $\hat{\sigma}_\eta$ будет малой.

6.4. Упражнения и лабораторный практикум

Упражнение 6.1.

Каков уровень значимости в проверке (сравнением математических ожиданий откликов) гипотезы адекватности имитационной модели? Выборка результатов наблюдений над системой: 5.5, 6.5, 7, результаты прогонов модели: 5, 6, 5.5, 5, 6.5, 6.5, 7.5, 5, 5.5, 5. Ответ 0.315.

Упражнение 6.2.

В антитетическом методе понижения дисперсии оценка коэффициента корреляции результатов параллельных расчётов оказалась равной – 0.7. Во сколько раз уменьшится погрешность имитационных расчётов? Ответ 2.58.

7. Применение имитационного моделирования к сравнению методов оценивания и анализу их точности

Для решения разнообразных задач (например, эконометрики) исследователи предлагают всё новые и новые модели и методы. Подчас проанализировать аналитически (дедуктивно) характеристики точности и прогностической силы этих методов бывает затруднительно. И здесь приходит на помощь ИМ.

Пример 8.1. Требуется найти оценку функции регрессии $y = \beta_0 + \beta_1 x$. Данные (наблюдения) имеют вид: $Y^n = \beta_0 + \beta_1 x^n + E^n$, где E^n – случайная ошибка (особенность) n -го наблюдения. Будем считать E^n *i.i.d.* $N(0, \sigma)$.

Допустим, эксперимент активный, т. е. значения x^n фактора исследователи могут выбирать сами.

Пусть есть два исследователя. Оба собираются оценивать вектор коэффициентов регрессии по методу наименьших квадратов (МНК):

$$\beta_{\text{best}} := (Z^T \cdot Z)^{-1} \cdot Z^T \cdot y, \quad (25)$$

$$\text{где } \beta_{\text{best}} := \begin{pmatrix} \beta_{\text{best } 0} \\ \beta_{\text{best } 1} \end{pmatrix}, \quad Z := \begin{pmatrix} 1 & x^1 \\ 1 & x^2 \\ \vdots & \vdots \\ 1 & x^N \end{pmatrix}, \quad y := \begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^N \end{pmatrix}.$$

Первый исследователь считает, что значения x^n фактора (матрицу плана Z) можно выбирать произвольно, второй же высказал догадку, что значения x^n лучше брать ортогонально к столбцу из единиц, т. е. $Z^{\langle 1 \rangle T} \cdot Z^{\langle 2 \rangle} = 0$, или

$$\sum_{n=1}^N x^n = 0, \text{ что, якобы, даст «более точные» оценки } \beta_{\text{best}}.$$

Как проверить, чей метод лучше?

Идея ИМ: закладываются известные истинные параметры β_0 и β_1 , рассчитываются для каждого исследователя «наблюдённые» значения y^1, y^2, \dots, y^N по формуле $Y^n = \beta_0 + \beta_1 x^n + E^n$. В ней N значений факторов, выбраны у них по-разному, N сгенерированных ошибок, одинаковых для обоих методов,. Затем по этим двум наборам оцениваются β_{best} по (25), и видно, что даёт более близкие к истине оценки (лучше)!

Чтобы «набрать статистику», проведём M прогонов:

$$\underline{N} := 5 \quad \underline{M} := 10 \text{ – число прогонов, } \beta := \begin{pmatrix} 1.5 \\ 0.5 \end{pmatrix}$$

$$z1 := \begin{pmatrix} 1 & 8 \\ 1 & 10 \\ 1 & 5 \\ 1 & 8 \\ 1 & 9 \end{pmatrix},$$

$$z2 := \begin{pmatrix} 1 & 8 \\ 1 & 10 \\ 1 & -8 \\ 1 & -5 \\ 1 & -5 \end{pmatrix}$$

$$\sigma := 1$$

$$\underline{m} := 1..M$$

$$\underline{e}^{(m)} := \text{rnorm}(N, 0, \sigma)$$

$$y1^{(m)} := z1 \cdot \beta + e^{(m)}$$

$$y2^{(m)} := z2 \cdot \beta + e^{(m)}$$

В каждом прогоне оцениваются параметры по N «наблюдениям»:

$$\beta_{\text{best}1}^{(m)} := (z1^T \cdot z1)^{-1} \cdot z1^T \cdot y1^{(m)}$$

$$\beta_{\text{best}2}^{(m)} := (z2^T \cdot z2)^{-1} \cdot z2^T \cdot y2^{(m)}$$

$$\beta_{\text{best}1} =$$

	1	2	3
1	1.583	2.826	3.159
2	0.384	0.421	0.352

$$\beta_{\text{best}2} =$$

	1	2	3
1	0.654	2.196	1.978
2	0.524	0.427	0.54

Поскольку каждый прогон даёт независимую реализацию случайной величины β_{best} , то среднее по всем прогонам даёт несмещённую и состоятельную (по

M) оценку математического ожидания $Best$, а $Stdev()$ даёт состоятельную и асимптотически несмещённую оценку стандартного отклонения $Best$. В частности, для оценки углового коэффициента (который, как мы знаем $\beta_1 = 0.5$):

$$\text{mean}\left[\left(\beta_{\text{best1}}^T\right)^{\langle 2 \rangle}\right] = 0.375$$

$$\text{mean}\left[\left(\beta_{\text{best2}}^T\right)^{\langle 2 \rangle}\right] = 0.476$$

$$\text{Stdev}\left[\left(\beta_{\text{best1}}^T\right)^{\langle 2 \rangle}\right] = 0.157$$

$$\text{Stdev}\left[\left(\beta_{\text{best2}}^T\right)^{\langle 2 \rangle}\right] = 0.057$$

Видно, что второй метод даёт более близкие к истине (0.5) результаты (с меньшим рассеиванием). На этом сравнение данных, якобы теоретически не исследованных методов, можно считать законченным.

Зная теорию, теперь можно её подтвердить:

1. Увеличивая M , убеждаемся, что математическое ожидание оценки коэффициента в обоих методах равно 0.5, т. е. они дают несмещённые оценки.

2. Увеличивая M , убеждаемся, что σ оценки коэффициента примерно = 0.271 и 0.059 (найжены при $M = 1000$).

Теория даёт:

$$\sigma_0 \cdot \sqrt{\left[(z_1^T \cdot z_1)^{-1}\right]_{2,2}} = 0.267$$

$$\sigma_0 \cdot \sqrt{\left[(z_2^T \cdot z_2)^{-1}\right]_{2,2}} = 0.06$$

что эквивалентно в последнем случае: $\frac{\sigma_0}{\sqrt{8^2 + 10^2 + (-8)^2 + (-5)^2 + (-5)^2}} = 0.06$.

7.1. Упражнения и лабораторный практикум

Упражнение 7.1.

Проверить имитационным моделированием целесообразность выбора ортогональной матрицы плана во множественном регрессионном анализе. [6]

Упражнение 7.2.

Проверить имитационным моделированием для какого-либо критерия, что SL имеет равномерное распределение при справедливости H_0 . Для этого в *MathCAD* запрограммировать, например, критерий согласия Колмогорова или хи-квадрат. Многократно на вход критерия генерировать выборки с распределением, совпадающим с теоретическим, заложенным в критерий. Построить гистограмму получающихся sl . [15], Т.1, с.228 – 235.

8. Задания для самостоятельной работы

Следующие задания могут служить темами курсовых работ. Каждая курсовая работа должна включать постановку задачи, детальную блок-схему программы, описание параметров и переменных программы, текст программы на соответствующем языке. Для многокомпонентных систем привести соображения по выбору метода организации квазипараллелизма и протяжки системного времени. Расчёты выполнять в *MathCAD* или специализированных пакетах моделирования. Провести верификацию модели и, по возможности, проверку адекватности. Обсудить полученные результаты, сделать выводы.

1. Проверка качества генератора псевдослучайных чисел. Применение критерия согласия хи-квадрат. Генератор *runif* (1000, 0, 1).

2. Проверка качества генератора псевдослучайных чисел. Применение критерия согласия Колмогорова. Генератор *rnorm* (1000, 0, 1).

3. Проверка качества генератора псевдослучайных чисел. Применение критерия случайности. Генератор *rexp* (1000, 1).

4. Вычисление интегралов методом Монте-Карло. Вычислить интеграл от $\sin(x)$ на $[0, \pi/2]$ с точностью 0.5% (с вероятностью 95%) с понижением дисперсии, используя псевдослучайные числа с плотностью распределения $\sim x$.

5. Вычисление интегралов методом Монте-Карло. Вычислить интеграл от $\sqrt{\sin(x)}$ на $[0, \pi/2]$ с точностью 0.5% (с вероятностью 95%) с понижением дисперсии, используя псевдослучайные числа с плотностью распределения $\sim \sqrt{x}$.

6. Вычисление интегралов методом Монте-Карло. Вычислить интеграл от $\exp(x)$ на $[0, 1]$ с точностью 0.5% (с вероятностью 95%) с понижением дисперсии, используя псевдослучайные числа с плотностью распределения $\sim 1+x$.

7. Имитационное моделирование надёжности сложных изделий. Изделие состоит из четырёх последовательно соединённых звеньев. 1, 2 и 4 звено имеют $MTBF$ 700, 500 и 600 часов. 3-е звено состоит из двух параллельно включённых для повышения надёжности элементов с $MTBF$ 400 и 300 часов. Найти $MTBF$ всего изделия с точностью 5% (с вероятностью 95%). Считать времена наработки на отказ распределёнными экспоненциально. [16], с.38 – 44.

8. Моделирование процесса случайного блуждания ξ_n – удаление вправо (в шагах) от исходной точки после n шагов. Шаги делаются только вправо или влево, одинаковой длины, независимо один от других. Вероятность шага вправо $= p$ (например, 0.5). Строить графики реализаций процесса ξ_n ($1 \leq n \leq N$). Оценить $M\xi_N$, $M|\xi_N|$ и $\sqrt{M(\xi_N)^2}$, где N , например, 100. [10], с. 85 – 86.

9. Имитационное моделирование инвестиционных рисков. Обобщение примера 4.3. Сделать нормы дисконта различными по годам: 0.2, 0.15, 0.1, 0.05, 0. Найти математическое ожидание NPV с точностью 0.5% (с вероятностью 95%). Оценить риск долей наблюдаемых NPV , меньших нуля. Оценить коэффициент корреляции NPV и средних за время проекта переменных затрат VAV . Оценить линейную функцию регрессии NPV на VAV и построить её график на фоне корреляционного поля (диаграммы рассеивания).

10. Имитационное моделирование инвестиционных рисков. Обобщение примера 4.3. Сделать ставку налога на прибыль случайной $\sim U(0.3, 0.9)$, найти то же, что в предыдущем задании.

11. Имитационное моделирование инвестиционных рисков. Обобщение примера 4.3.

Показатели	Сценарии		
	Наихудший $P = 0.25$	Наилучший $P = 0.25$	Вероятный $P = 0.5$
Q	150	300	200
P	40	55	50
V	35	25	30

По сценариям рассчитать средние и дисперсии показателей и использовать их в нормальной аппроксимации распределения соответствующих величин, считая их независимыми. Найти то же, что в предыдущем задании.

12. Имитационное моделирование инвестиционных рисков. Обобщение примера 4.3. Считать Q , P , V независимыми дискретными случайными величинами с законами распределения – как таблицы в предыдущей теме. Найти то же, что в предыдущем задании.

13. Имитационное моделирование инвестиционных рисков. Обобщение примера 4.3. По сценариям из задания № 11 найти то же, что в предыдущем задании.

14. Имитационное моделирование инвестиционных рисков. Обобщение примера 4.3. В условиях задания № 11 считать P и V зависимыми с $\rho = 0.5$. Найти то же, что в предыдущем задании. Сравнить оценки коэффициентов корреляции NPV и VAV при $\rho = 0$ и 0.5 . Дать экономическое объяснение.

15. Моделирование управления запасами. Начальный запас – 2000 шт., еженедельный спрос $\sim N(500, 10)$, стоимость хранения штуки запаса на конец недели –

0.5 у. е., каждая единица неудовлетворённого спроса оценивается потерями 20 у. е. Если запас на конец недели опускается ниже 1500 шт., то делается заявка на 1000 шт., на выполнение которой (к началу недели) может потребоваться случайное число недель:

Недели	1	2	3	4
Вероятность	0.2	0.5	0.25	0.05

Пока предыдущая заявка не выполнена, новая не может быть подана. Стоимость подачи заявки – 50 у. е. Оценить ожидаемую стоимость в неделю. Определить с точностью до 10 шт. оптимальный размер заявки и соответствующую стоимость в неделю.

16. Задача об управлении персоналом. Компании «Три звездочки» необходимо решить задачу об оптимальном количестве гидов, необходимых для проведения экскурсий. Ежедневная заработная плата постоянных гидов составляет 32 песо. Если на всех клиентов гидов не хватает, то можно нанять сторонних специалистов. Стоимость их использования составит 90 песо в день. Может случиться, что гидов всё равно не хватит, тогда компания терпит убытки в размере 500 песо за каждую сорванную экскурсию.

Было произведено 100 наблюдений за спросом на гидов. См. табл.

Спрос	4	5	6	7	8	9	10	11	12	13	14	15	16
Частота	0.02	0.04	0.04	0.08	0.1	0.12	0.12	0.12	0.1	0.1	0.08	0.06	0.02

Данные по предложению дополнительных гидов в выходные и в будни различаются:

Предложение в будни	0	1	2	3	4	5	6	7
---------------------	---	---	---	---	---	---	---	---

Ча- стота	0.03	0.09	0.15	0.2	0.22	0.14	0.1	0.07
--------------	------	------	------	-----	------	------	-----	------

Предложение по вос- кресеньям	0	1	2	3	4	5	6	7
Частота	0.14	0.26	0.28	0.18	0.08	0.04	0.02	0

Необходимо определить оптимальное число постоянных гидов так, чтобы минимизировать издержки, возникающие при работе компании. Каковы эти минимальные издержки в неделю? [17], с.188 – 196.

17. Моделирование процесса обслуживания клиентов в банке. В банке 2 кассы, приход клиентов описывается пуассоновским потоком с интенсивностью $\lambda = 60$ чел./час. Время обслуживания клиентов – экспоненциальное с интенсивностью $\mu = 45$ чел./час. Вошедший клиент присоединяется к кратчайшей очереди. Проимитировать работу банка в течение 80 часов.

18. Моделирование работы парикмахерской с двумя креслами обслуживания. Интервалы времени между клиентами имеют равномерное распределение от 5 до 12 минут, времена обслуживания 1-м и 2-м парикмахерами имеют равномерное распределение от 12 до 20 минут. Проимитировать работу парикмахерской в течение 8 часов.

19. Моделирование двух стратегий обслуживания на автозаправочной станции. АЗС имеет 2 бензоколонки, автомобили прибывают через каждые $\xi \sim U(30, 70)$ сек., а время заправки их бензином составляет $\eta \sim U(50, 130)$ сек. Можно предложить две стратегии обслуживания прибывающих автомобилей: 1) автомобили выбирают в случайном порядке одну из 2-х очередей;

2) автомобили образуют одну очередь к двум колонкам. Выбрать лучшую стратегию.

20. Планирование бизнеса методами ИМ. Поток автомобилей на круглосуточной моечной станции является пуассоновским с интенсивностью 9 (1/час), время мойки распределено экспоненциально, среднее время – 5 мин. Если клиенты не застают свободного места на стоянке, они уезжают. Упущенная прибыль – 20 руб./авто. Аренда земли под стоянку 1 автомобиля – 20000 руб. в год. Сколько следует иметь мест на стоянке? Каковы минимальные потери (руб./сутки)?

21. Моделирование процесса функционирования промышленного робота. Поступающие к роботу изделия образуют пуассоновский поток с интенсивностью 0.02 изд./сек, а время обработки изделия составляет 45 сек. Вероятность безотказной работы робота за время обработки изделия равна 0.95. Если робот отказывает, то происходит его восстановление по экспоненциальному закону с математическим ожиданием 70 сек. После восстановления робота обработка изделия продолжается, причём фиксируется общее число отказов робота. Проимитировать обработку 10000 изделий.

22. Решение задачи коммивояжёра методом Монте-Карло. Найти кратчайший (по возможности) маршрут, включающий все города и начинающийся и заканчивающийся в 1-м городе, и его длину. [14], с.460.

Города	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	9	9	1	9	4	1	2	1	1	9	4	8	4	1	1	9	7	6	6
2		1	9	7	6	2	5	1	9	7	2	6	3	1	9	5	1	1	1
3			9	5	5	8	1	9	8	1	3	2	1	9	4	9	5	5	7
4				5	8	9	4	2	1	4	5	9	4	1	2	3	4	5	8
5					5	5	6	4	7	8	9	1	3	9	5	4	8	3	9
6						8	9	7	6	4	1	2	2	3	8	5	1	1	5
7							5	7	7	8	1	5	9	2	9	5	2	3	6
8								5	8	9	8	6	4	3	8	5	6	5	7

9									4	3	8	5	1	9	2	6	6	7	2
10										2	3	1	1	5	4	4	6	5	2
11											5	7	3	8	9	4	2	1	8
12												5	4	2	1	6	3	5	5
13													7	3	2	8	6	5	3
14														7	9	3	2	1	4
15															5	9	8	2	1
16																4	6	7	9
17																	9	8	5
18																		5	3
19																			3

23. Моделирование случайных процессов. Сгенерировать реализацию процесса $\eta_n = c_0\xi_n + c_1\xi_{n-1} + c_2\xi_{n-2}$, ($n = 1, 2, \dots, N$), где ξ_n – последовательность независимых случайных величин $\sim U(-1, 1)$; c_0, c_1, c_2 – постоянные коэффициенты, $c_0 + c_1 + c_2 = 1$.

Построить график этой реализации (временного ряда). Вычислить оценки математического ожидания и ковариационной функции процесса η_n : $\hat{m} = (y_1 + y_2 + \dots + y_N)/N$ и $\hat{K}(s) = [\sum_{n=1}^{(N-s)} (y_n - \hat{m})(y_{n+s} - \hat{m})]/(N-s)$, $s = 0, 1, 2$, и сравнить их с теоретическими значениями.[18].

24. Определение цены опциона методом ИМ. При начальной стоимости актива $s_0 = 1000$ и цене исполнения $x = 1100$ найти цену одногодичного опциона (250 торговых дней) с точностью 5% (с вероятностью 0.95). Для однодневной ставки непрерывного наращения R собраны 2250 наблюдений, которые распределены по интервалам $(-0.0325, -0.03), (-0.03, -0.0275), \dots,$

$(0.03, 0.0325)$ так:

5,6,7,8,13,23,38,50,95,144,185,213,241,246,262,239,185,134,77,26,19,15,9,5,4,1.

25. Проверка качества генератора псевдослучайных чисел. Определить экспериментально период и длину отрезка аperiodичности какого-либо генератора псевдослучайных чисел. [3], с.127.

26. Планирование бизнеса методами ИМ. Пресс-конференция для K журналистов проходит в зале, где имеется K телефонов. По окончании пресс-конференции каждый журналист бросается к телефону. Предположим, что все K телефонов равнодоступны любому журналисту и что он выбирает один из них случайным образом, независимо от остальных журналистов. Если несколько журналистов выбрали один и тот же телефон, то он закрепляется лишь за одним из них, а остальные вынуждены выбирать среди оставшихся телефонов до тех пор, пока каждый не найдёт себе телефона.

Выяснить, сколько в среднем попыток (в зависимости от K) должно быть сделано до того, как каждому журналисту достанется телефон. В частности, найти математическое ожидание числа попыток при $K = 10$ с погрешностью 0.2 % (с вероятностью 0.95). Как повлияло бы (при числе журналистов 10) появление i дополнительных телефонов, $i = 1, 2, 3$?

27. Использование пакета *ARENA v.2.2* в ИМ СМО (для верификации повторить пример 5.3).

28. Использование пакета *GPSS* в ИМ СМО (одно из заданий 15 – 21). [19, 20].

29. Использование пакета *EXTEND* в ИМ СМО (для верификации повторить пример 5.3).

30. Понижение дисперсии антитетическим методом в задаче моделирования инвестиционных рисков. Повторить с понижением дисперсии пример 4.3. Найти математическое ожидание NPV с точностью 0.05% (с вероятностью 99.7%).

31. Планирование бизнеса методами ИМ. Время между обращениями механиков за инструментами на большом заводе распределено экспоненциально со средним 10 мин. Время выполнения запроса нормально со средним 7.5 мин. и стандартным отклонением 2 мин. Время простоя механика стоит 0.2 руб./мин., а кладовщика – 0.15 руб./мин. Сколько должно быть кладовщиков, чтобы суммарные потери от простоя механиков и кладовщиков были минимальны? Каков этот минимум (в час)?

32. Планирование бизнеса методами ИМ. В ремонтной мастерской ломается в среднем 0.6 резцов в день. Среднее время выполнения заявки на поставку резцов 6 дней. (Резцы поставляют раз в день к началу рабочего дня). Число поломок и время выполнения заявки имеют распределения Пуассона. Резцы заказываются в начале рабочего дня, когда 3 резца (или больше) должны быть заменены. Более одной заявки в стадии выполнения не допускается. Какое минимальное количество резцов должно быть в мастерской, чтобы, по крайней мере, 1 исправный резец был в 95% рабочих дней?

33. Моделирование случайных векторов. Распределение $p_{\vec{x}}(\vec{x}) = \begin{cases} ax_2, \vec{x} \in D \\ 0, \vec{x} \notin D \end{cases}$,

D – треугольник с вершинами (0,0), (0,1), (1,0). [3], с.140 – 141.

34. Моделирование случайных векторов. Оценив необходимые параметры по юношам (девушкам) своей студенческой группы, сгенерировать выборку 100 юношей (девушек). Интересующие параметры – рост, вес. [3], с.140 – 142.

35. Применение ИМ к сравнению методов оценивания и анализу их точности. Пусть система имеет вид $Y = \beta_1 X_1^2 + \beta_2 X_2^2$, где X_1 и X_2 – независимые $\sim N(0,1)$ случайные величины. Строим линейную модель без константы, МНК-оценки ко-

торой находятся по известной формуле. Показать имитационным моделированием, что с увеличением числа наблюдений оценки коэффициентов модели стремятся к 0 (несостоятельность!). Подтвердить этот результат дедуктивно.

36. Применение ИМ к сравнению методов оценивания и анализу их точности. Имитационным моделированием найти мощность какого-либо критерия. По возможности сравнить с точным значением (например, для t – критерия проверки гипотезы о математическом ожидании).

37. Методом статистического моделирования решить следующую задачу. Проводится 10 независимых выстрелов по мишени, причём вероятность попадания при одном выстреле равна p . Требуется определить вероятность того, что число попаданий в мишень будет чётным, т. е. 0, 2, 4, 6, 8, 10. [3], с.113.

38. Антитетический метод понижения дисперсии в вычислении интеграла обычным методом Монте-Карло (используя равномерно распределённые псевдослучайные числа). Вычислить интеграл от $\sqrt{\sin(x)}$ на $[0, \pi/2]$ с точностью 0.5% (с вероятностью 95%).

39. Оценка погрешности результирующего показателя имитации из-за различия *seed*. Пусть известно, что на систему воздействует входная случайная величина $L \sim U(-1, 2)$ и влияет внешняя среда, случайная величина $F \sim N(2, 1)$. Нас интересует выходная величина $Y = \sqrt{(1 + e^{-L})^2 + (1 - e^{-F})^2}$. Найти с точностью 1% (с вероятностью 95%) максимальное значение Y .

40. Имитационное моделирование кредитных рисков. Рассмотреть кредитный портфель, состоящий из большого займа «лидера» и нескольких десятков меньших займов, дефолты по которым одинаково коррелированы с дефолтом «лидера». Задать реальные (правдоподобные) параметры. Исследовать ожидаемые

потери, случайные потери, экономичный капитал в различных случаях. Сделать выводы [22].

Библиографический список

Основная литература

1. Емельянов А. А. Имитационное моделирование экономических процессов : учеб. пособие для студентов, обучающихся по специальности «Прикладная информатика» / А. А. Емельянов, Е. А. Власова, Р. В. Дума ; под ред. А. А. Емельянова. – М. : Финансы и статистика, 2004.
2. Лоу А. М. Имитационное моделирование / А. М. Лоу, В. Д. Кельтон ; перевод с англ. А. Куленко под ред. В. Томашевского. – 3-е изд. – М. ; СПб. ; Нижний Новгород [и др.] : Питер, 2004.
3. Советов Б. Я. Моделирование систем : учеб. для вузов / Б. Я. Советов, С. А. Яковлев. – 3-е изд., перераб. и доп. – М. : Высш. шк., 2001.

Дополнительная литература

4. Максимей И. В. Имитационное моделирование на ЭВМ / И. В. Максимей. – М. : Радио и связь, 1988.
5. Лукасевич И. Я. Анализ финансовых операций / И. Я. Лукасевич. – М. : ЮНИТИ, 1998.
6. Ивченко Г. И. Математическая статистика / Г. И. Ивченко, Ю. И. Медведев. – М. : Высш. шк., 1984.
7. Большев Л. Н. Таблицы математической статистики /Л. Н. Большев, Н. В. Смирнов. – М. : Наука, 1983.
8. Исследование операций в экономике / ред. Кремер Н. Ш. / М. : ЮНИТИ, 1997.
9. Айвазян С. А. Прикладная статистика и основы эконометрики / С. А. Айвазян, В. С. Мхитарян. – М. : ЮНИТИ, 1998.
10. Харин Ю. С. Основы имитационного и статистического моделирования / Ю. С. Харин, В. И. Малюгин, В. П. Кирлица и др. – Минск : Дизайн ПРО, 1997.

11. Уотшем Т. Дж. Количественные методы в финансах / Т. Дж. Уотшем, К. Паррамоу. – М. : ЮНИТИ, 1999.
12. Плис А. И. *Mathcad 2000*. Математический практикум для экономистов и инженеров : учеб. пособие / А. И. Плис, Н. А. Сливина. – М. : Финансы и статистика, 2000.
13. Клейнен Дж. Статистические методы в имитационном моделировании : в 2 т. [пер. с англ.]. / Дж. Клейнен. – Серия: Математико-статистические методы за рубежом. – М. : Статистика, 1978.
14. Коршунов Ю. М. Математические основы кибернетики / Ю. М. Коршунов. – М. : Энергоатомиздат, 1987.
15. Справочник по прикладной статистике: в 2 т. / Ред. Э. Ллойд, У. Ледерман. – М. : Финансы и статистика, 1990.
16. Соболев И. М. Метод Монте-Карло / И. М. Соболев. – М. : Наука, 1972.
17. Кофман А. Займёмся исследованием операций / А. Кофман, Р. Фор. – М. : Мир, 1966.
18. Чистяков В. П. Курс теории вероятностей / В. П. Чистяков. – М. : Наука, 1982.
19. Пакет программ GPSS. Режим доступа:
<http://www.minutemansoftware.com/downloads/student.exe>
20. Кудрявцев Е. М. *GPSS World*. Основы имитационного моделирования различных систем / Е. М. Кудрявцев. – М. : ДМК Пресс, 2004.
21. Бородачёв С. М. Методы математической статистики : учебное пособие / С. М. Бородачёв. – Екатеринбург : УрФУ, 2012. – 127 с. ISBN 978–5–321–02218–4
22. *Bluhm, Christian. An Introduction to Credit Risk Modeling. Chapman & Hall, 2003.*
23. Yahav I., Shmueli G. On generating multivariate Poisson data in management science applications. <http://onlinelibrary.wiley.com/doi/10.1002/asmb.901/references>

Учебное издание

Бородачёв Сергей Михайлович

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ В ЭКОНОМИКЕ

Редактор *В. И. Новикова*

Компьютерная верстка *авторская*

Подписано в печать . Формат 60 x 84 1/16.

Бумага писчая. Плоская печать. Усл. печ. л.

Уч.-изд. л.. Тираж экз. Заказ

Редакционно-издательский отдел УрФУ

620002, Екатеринбург, ул. Мира, 19

rio@mail.ustu.ru

Ризография НИЧ УрФУ

620002, Екатеринбург, ул. Мира, 19